②

REPORT DOCUMENTATION PAGE

DTIC FILE COPY

Form Approved
OMB No. 0704-0188

AD-A226 809

1a. REPORT SECURITY CLASSIFICATION Unclassified		1b. RESTRICTIVE MARKINGS	
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE 1990		*unlimited*	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) R910023-F		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 90 09 23	
6a. NAME OF PERFORMING ORGANIZATION Scientific Research Associates	6b. OFFICE SYMBOL *(If applicable)* 8N189	7a. NAME OF MONITORING ORGANIZATION *Same as 8a*	
6c. ADDRESS (City, State, and ZIP Code) Nye Road Glastonbury, CT 06033		7b. ADDRESS (City, State, and ZIP Code) *Same as 8c*	
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Air Force of Scientific Research	8b. OFFICE SYMBOL *(If applicable)* ~~FQ8671~~ NE	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F49620-87-0055	
8c. ADDRESS (City, State, and ZIP Code) Bolling Air Force Base Washington, DC 20332		10. SOURCE OF FUNDING NUMBERS	

	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
	61102F	2306	B1	

11. TITLE (Include Security Classification)
Studying Quantum Phase-Based Electronics Devices

12. PERSONAL AUTHOR(S)
H. L. Grubin, M. Cahay, J. P. Kreskovsky

13a. TYPE OF REPORT Final	13b. TIME COVERED FROM 87/05/20 TO 90/06/14	14. DATE OF REPORT (Year, Month, Day) 90/08/15	15. PAGE COUNT

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Quantum Transport, Wigner Functions, Density Matrix Equations, Aharanov-Bohm Effect, Quantum Wells, Quantum Barriers, Resonant Tunneling Structures, Quantum Diffraction Effects

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A study was undertaken to examine, theoretically, quantum phase-based electronic device. The study was implemented through examination of the moments of the Wigner Distrubition Function, time dependent solutions to Schrodingers equation in two dimensions and solutions to the equation of motion of the Density Matrix. Solutions were obtained using numerical methods. A variety of problems were considered including simulation of resonant tunneling structures, electron diffraction through an aperature in a potential well, and examination of the Aharanov-Bohm effect.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT [X] UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code) 202-767-4931	22c. OFFICE SYMBOL NE

DD Form 1473, JUN 86
Previous editions are obsolete.

90 09 25 011

Contract No.: F49620-87-C-0055
SRA No.: R90-910023-F

Final Report
Period: 20 May 1987 - 14 June 1990

STUDYING QUANTUM PHASE-BASED ELECTRONIC DEVICES

H. L. Grubin, M. Cahay and J. P. Kreskovsky

Scientific Research Associates, Inc.
P. O. Box 1058
Glastonbury, CT 06033

Prepared for

Air Force Office of Scientific Research
Bolling Air Force Base, D.C.

August 1990

Scientific Research Associates, Inc.
P. O. Box 1058
Glastonbury, CT 06033

1.INTRODUCTION

Since the pioneering work of Tsu and Esaki [1], there has been growing interest in the double-barrier resonant tunneling two and three terminal devices [2-13]. Further, since the initial proposal of Datta et al [14], there has also been rising interest in the basic physics accompanying the Aharonov-Bohm effect in heterostructures. Indeed, major advances in material technology has enabled device scientists to conjecture about new device structures that both test and illustrate the basic fundamental quantum physics of few and many particle systems. For example issues of nonlocality, often relegated to esoterica, now find its way into discussions on transport in quantum devices. Nonlocality in classical physics is illustrated by the coulomb interaction that decreases as the square of the distance between particles. In quantum mechanics there are additional interactions that do not necessarily drop off with distance. The effect of nonlocality is represented in the quantum mechanical equations and in the boundary conditions to the system. A second issue involves dissipation. Schrodinger's equation is dissipationless, and if all transport in subsystems were governed by Schrodingers equation without interactions between the subsystems, all transport would be ballistic. Dissipation in quantum mechanics is treated by introducing additional systems, e.g. a phonon system, and allowing the additional systems to cause a transition between states of the original system.

Our ability to use quantum mechanics to describe physical phenomena in ultrasmall devices and to propose quantum phase based devices has been evolutionary. Through a coupling of experiment, theory and numerical simulation we are better able to understand how basic quantum mechanical processes affect device physics. An evolutionary parallel exists with the more standard semiconductor FETs that are currently being used in DOD and commercial applications. We digress for a moment to illustrate.

Originally, when FETs were discussed it was in terms of their DC characteristics; and when these devices were considered, e.g., for linear amplifier applications, the DC characteristics were used to obtain such small signal quantities as transconductance. Unfortunately the dc characteristics do not display the basic properties of the FET. They hide the influence of traps, charge particle instabilities, etc. To demonstrate this, nearly a decade ago, workers [15], currently at Scientific Research Associates, Inc., based on theoretical studies suggested that the large signal characterization of a device is significantly different than the dc characterization. Experimental work, carried out in collaboration with these workers confirmed this result. We display this result in figure 1. **The key point to notice about this figure is that the dc and transient characterization of the FET are dramatically different!** Indeed, in response to this dilemma, the MIMIC community is questioning whether the DC characterization of FETs has any relevance in the design of power FETs.

The situation with quantum phase based devices is **similar** to that of the FET modeling over ten years ago. With few exceptions, e.g., of some Wigner studies [16] most of the analysis is time-independent. This has introduced interpretive difficulties:

First: the dc studies, do not account for the peak to valley ratio of resonant tunneling devices.

Second: the dc studies do not adequately treat dissipation.

-1-

Third: the dc studies do not treat hysterisis in the current voltage characteristics, observed experimentally.

Fourth: the dc treatment cannot predict how the devices will be used in applications.

Fifth: the dc treatment cannot treat the time dependent nature of the boundary conditions that represent physical contacts,

Sixth: the studies associated with the A-B effect are all time independent and small signal, and the limits of the physics of this phenomena are unknown.

The above studies suffer from lack of incorporating the feature basic to quantum mechanical phenomena: **all quantum mechanical devices are time dependent.** Apart from dissipation, there are always reflections off boundaries, barriers, wells, imperfections and contacts. **What is needed is a full time dependent large signal numerical studies of quantum feature size devices.**

The purpose of this document is to summarize studies under Contract F49620-87-C-0055, the study of the physics of quantum based electronic devices, that is, devices whose only description is quantum mechanical.

2. SUMMARY OF STUDIES UNDER CONTRACT F49620-87-C-0055

During the current program a number of key steps were taken in the direction of developing a time dependent algorithm for examining quantum phase based devices. Some of these steps are summarized below. These studies included:

(1) implementation of a fully self-consistent (Poisson's equation is included) solution to the moments of the Wigner distribution function,

(2) implementation of a two dimensional Schrodinger equation solver for examining quantum diffraction phenomena, and

(3) initial implemention of a density matrix solver for examining quantum transport.

2a. MOMENTS OF THE WIGNER DISTRIBUTION FUNCTION

This study started from the equation of motion of the Wigner distribution function [17]. Assuming a specific form for the Wigner function a set of moment equations was derived. The moment equations are regarded as the quantum mechanical equivalent of the semiclassical moments of the Boltzmann transport equation, and are designed to enable broad use by the scientific community in examining quantum device physics. Three equations were derived: the continuity equation, the momentum balance equation, and the energy balance equation. Implementing the first two equations, transport in resonant tunnelling structure was studied. The phenomena of resonance was observed. It remains to incorporate the energy balance equation to obtain the correct voltage levels for resonance. Two paper discussing the moment equations were published. They are included as Appendix A and B of this document.

2b. TWO DIMENSIONAL TIME DEPENDENT
SOLUTION TO SCHRODINGERS EQUATION

Most studies of the switching response of quantum phase based devices are one dimensional. In practice electrons are injected through 3D contacts and tunnelling through the quantum well of resonant tunnelling diodes is characterized by at least two dimensional transport. This calls for a two dimensional transient analysis of quantum transport. Presently the two dimensional analysis is beyond the scope of the density matrix discussion. Instead preliminary steps were taken to address this problem by examining the time evolution of a wavepacket through a multidimensional slit.

Briefly, we studied the diffraction of a two-dimensional Gaussian wavepacket through a rectangular aperture in a finite potential well (a single slit experiment). The simulations show a small fraction of the wavepacket tunneling through the potential barrier. But the dominant transport shows a near field diffraction pattern behind the slit for a small time duration (2ps). At a later time, no far field diffraction pattern is seen.

A copy of a paper published on this matter is included as Appendix C.

2c. AHARONOV-BOHM EFFECT

Most recently we were able to predict the existence of two different sets of conductance minima in the conductance oscillations of a one dimensional ring due to the electrostatic A-B effect.

A copy of a paper summarizing this work is included as Appendix D.

2d. THE DENSITY MATRIX

The technique of choice is to solve the equation of motion of the density matrix in the coordinate representation. The solution to the equation of motion of the density matrix is:

1. Time dependent,

2. More general that the Tsu-Easki formulation, which apart from severe analytical approximations, treats the current carrying states as stationary,

3. Capable of treating dissipation.

Under the present study we have already solved this equation for several different configurations and boundaries. While solutions to the problem of quantum transport through the density matrix in the coordinate representation has been proposed earlier by Frensley [18], **Scientific Research Associates, Inc. (SRA) is the first group that has been successful in obtaining solutions.** The resulting insight into the physics, garnered from these studies as we discuss below, places SRA in a unique position to interpret present experimental results. There are, of course, other time dependent methods that have been used to determine the operation of devices [16]. One mentioned above is the Wigner distribution approach. We have implemented the moments of the Wigner function in this aspect. Others have attempted to solve the full time dependent Wigner function. While the

density matrix approach used by the those at SRA and the Wigner approach are mathematically equivalent, the present state of *numerical* development *does not render the* Wigner and density matrix approach to be equivalent, and the wavelike nature of the **time dependent** solution as obtained from the Wigner approach is not always apparent.

The structures that we discuss fall under the category of "open structures" [18], which can exchange particles with its surrounding, and which mathematically expresses this interaction in terms of boundary conditions. The phenomena we are interested in will be with systems that are far from equilibrium.

In describing this nonequilibrium phenomena, the issues of interest include the means of describing the phenomena. This is nontrivial. For example, one of the main difficulties in trying to understand the precise changes implied by quantum mechanics lies in the formalism itself. It is very different from that of classical mechanics. In past studies under this contract we have adapted the approach of Bohm in describing quantum mechanics in a language that is closer to that of classical mechanics. This is achieved by writing the wave function of the system of interest in the form:

(1) $\psi = R \exp[iS/\hbar]$

By assuming that the wavefunction satisfies Schrodinger's equation we obtain two real equations, one of which has the form of a classical equation **supplemented by an additional term called the quantum potential**, and a second equation describing the conservation of probability. To understand these equations it is necessary to assume that there is a microscopic reality in which particles have both position and momentum, although these cannot be measured simultaneously. The solutions of these quantum mechanical equations of motion gives rise to an **ensemble** of individual particle trajectories arising from various initial conditions. If the distribution of initial conditions agrees with that calculated from the initial wave function, then the ensemble will give rise to the expected probability distributions found in experiment. This approach often suggests the possibility of introducing a quantum Monte Carlo formulation.

In previous studies at SRA, the description of the ensemble of individual particles was treated either through implementation of Wigner distribution procedures or through implementation of the density matrix in the coordinate representation. Since the Wigner distribution function and the density matrix in the coordinate representation are related through a direct integral transformation they contain the same physics. The advantages of using one against the other lies in the computational efficacy with which the underlying physics is exposed. It is anticipated that in the future, the choice will depend on personal preference. Under the present Contract both were used, although differently.

The density matrix or equivalently the Wigner function is the tool required to transition between measurements and theory for multiparticle systems such as tunneling devices, A-B devices etc. They represent the essence of mixed states and are at the basis of all theoretical descriptions of measurement. We describe these briefly below.

MIXED STATES AND THE DENSITY MATRIX

In general the system that is being studied is in a mixed state, a state to which a wave

function cannot be assigned. Mixed states are described as a collection of pure states $|i>$, where each pure state is assigned a weight $W(i)$. The weights are real and by convention they satisfy the condition:

$$(2) \qquad \Sigma_i W(i) = 1$$

A discussion of measurement centers about expectation values. For example, the expectation value of momentum, **p**, in the pure state $|i>$ is:

$$(3) \qquad <p(i)> = <i|p|i>$$

In the system in which each of the pure states is assigned the relative weight $W(i)$, the measurement is associated with an **ensemble** average, which in turn is related to the expectation value through the following expression:

$$(4) \qquad <<p>> = \Sigma_i W(i) <p(i)> = \Sigma_i W(i) <i|p|i>$$

The task of theory is to calculate the ensemble expectation values. At this point we need to do some arithmetic. We invoke the standard expansion and express the pure state $|i>$ in terms of a complete set of linearly independent eigenkets, $|m>$:

$$(5) \qquad |i> = \Sigma_m A_{m,i} |m>$$

In terms of these eigenkets the ensemble average of momentum is **rewritten** as:

$$(6) \qquad <<p>> = \Sigma_{i,m,n} W(i) A^*_{n,i} A_{m,i} <n|p|m>$$

The introduction of the density matrix is a compact way of writing the above expression of a measurement of the ensemble average value of momentum. We define the elements of the density matrix as:

$$(7) \qquad \rho_{m,n} = \Sigma_i W(i) A^*_{n,i} A_{m,i}$$

The density matrix with elements given by equation (7) is referred to below as the **density matrix in the 'state' representation**, to distinguish it from the density matrix in the 'coordinate' representation discussed below. The ensemble average of momentum in the state representation is:

$$(8) \qquad <<p>> = \Sigma \rho_{m,n} <n|p|m>$$

Thus if the eigenkets $|m>$ are known, the task of calculating the ensemble average expectation value involves calculating $\rho_{m,n}$. We will not discuss the properties of the density matrix in the state representation, identified in equation (7) as this is discussed in most textbooks. Rather, in the discussion that follows we will concentrate on the 'reduced' density matrix in the 'coordinate' representation. Our description will be within the framework of the single electron picture.

The density matrix in the coordinate representation is obtained from the density matrix in the state representation through the following prescriptive (we confine ourselves to one space dimension): If $\psi_m(x,t)$ is the Schrodinger wave function associated with the eigenstate

$|m>$, the density matrix in the coordinate representation is defined as:

(9a)
$$\rho(x,x',t) = \Sigma \rho_{m,n} \psi_m(x,t) \psi^*_n(x',t)$$

With regard to equation (9a) we are faced with the following choices in finding the density matrix in the coordinate representation:

(1) We can solve the differential equation for $\rho_{m,n}$, and Schrodinger's equation for the wavefunctions and then obtain the density matrix in the coordinate representation, or

(2) We can solve for the density matrix in the coordinate representation directly, and **ignore the intermediate Schrodinger equation.**

In the proposed program because we are dealing with a system of particles we choose the solve for the density matrix directly, and intend to develop a language that is suitable for the density matrix.

We point out that many discussions of the density matrix in the coordinate representation takeoff with the assumption that a matrix transformation can be found that diagonalizes the density matrix in the 'state' representation. In this case equation (9a) becomes:

(9b)
$$\rho(x,x',t) = \Sigma \rho_{m,m} \psi_m(x,t) \psi^*_m(x',t)$$

The assumption leading to equation (9b) is not restrictive. While we often invoke this assumption for initial conditions, the solutions to the problem are not dependent upon it.

The density matrix in the coordinate representation satisfies the equation of motion:

(10)
$$\partial \rho(x,x',t)/\partial t$$
$$= (i\hbar/2m)[\partial^2/\partial x^2 - \partial^2/\partial x'^2]\rho(x,x',t)$$
$$- (i/\hbar)[V(x,t) - V(x',t)]\rho(x,x',t)$$

We point out that the Wigner function is the Fourier transform of the density matrix. In one dimension:

(11)
$$f_w(p,x) = [1/(2\pi\hbar)]\int d\eta \rho(x+\eta/2, x-\eta/2)\exp[-ip\eta/\hbar]$$

Equation (10) is the differential equation of interest. It describes the temporal and spatial evolution of an ensemble of carriers in the system under investigation. It is the multiparticle equivalent of Schrodinger's equation, and reduces to the latter when $\rho(x,x't) = \psi(x,t)\psi^*(x',t)$. Rather than seeking solutions to Schrodingers equation for the wave function, we seek solutions to the equation of motion of the density matrix. It is discussed in detail as it is unfamiliar to many.

Solutions to equation (10) or the equivalent Wigner function are required for determining transport in quantum phase based devices. We point out that the differential equation for

the density matrix in the coordinate representation is a two dimensional differential equation; **yet it describes only one dimensional physics!** That is, there are two independent variables x and x'. But recalling that the density matrix is the Fourier transform of the Wigner function, which is also a function of two variables, x and p, we see that we are not adding additional concepts that were not in place before. The issue is interpretation.

The density matrix in the coordinate representation, with the exception of the diagonal elements, is not an observable. Rather, it appears in the integrand in a two dimensional integral in which the ensemble average of an operator is calculated. Thus, it appears to have some of the properties of a correlation function. For example if the wavefunction represented in equation (9) were the hydrogenic wavefunctions, then the density matrix would be strong only in the region where there was strong overlap in the wavefunctions. The same situation occurs here. In the following we will be imposing boundary conditions on the density matrix and therefore on the 'overlap' at the boundaries. These boundary conditions would be the mathematical equivalent of constraining the values of position and momentum on the Wigner function.

While the numerical problem discussed above is a two dimensional problem **the diagonal components of the density matrix are the observables and represent the ensemble average of the density of the system.** In this section we will illustrate the calculations of the diagonal components, extracting these results from the more general two dimensional solution discussed in detail below. We must point out some features of the solutions that are quantum based and have been known and discussed for well over 50 years. First the Wigner function which displays the closest connection to the Boltzmann distribution function and is sometimes referred to as a quantum probability distribution function, **can be negative!** Similarly, **the density matrix can be negative.** However, the density matrix is subject to the following constraints: $\text{Re}\rho(x,x') = \text{Re}\rho(x',x)$; $\text{Im}\rho(x,x') = -\text{Im}\rho(x',x)$. **Thus the density matrix along the diagonal is always positive,** as is required by the physics. The off-diagonal elements which control the values of the density matrix on the diagonal, can be negative. From the density matrix, we can obtain the ensemble expectation values of all observables. In the discussion that follows we will be interested in three observables: **density, current and energy.** We note that expectation values of operators in the coordinate representation and within the Wigner formulation are written, respectively, as (we display this for the momentum operator):

(12) $<<\mathbf{p}>> = \int dx\,dx'\, <x'|\mathbf{p}|x> \rho(x,x',t)$

(13) $<<\mathbf{p}>> = \int dx\,dp\, f_W(x,p,t)\,\mathbf{p}$

All of the situations of interest involving a detailed comparison of experiment with theory, as well as a prediction of the operation of devices and their dependence on material parameters require that the solutions obtained be coupled to Poisson's equation and that they be obtained at finite as well as zero bias. For one aspect of the study, namely that associated with the moments of the Wigner equation, this was done. The studies with the density matrix are new. Poisson's equation has not yet been coupled to the density matrix equation, nor are the results at finite bias. All of this is **currently** taking place, and will be continued during the proposed follow-on study.

In the next section we extract part of the full density matrix solution and display only the

solutions along the diagonal. A discussion of the full density matrix solution occupies the bulk of this document and is supplemented with a number of analytical discussions. Further, because the initial calculation are for zero bias, although one uniform density solution was for a finite bias, **the time dependent solutions are a consequence of a strong perturbation of a steady state solution.**

THE DENSITY DISTRIBUTION AS OBTAINED FROM
THE DENSITY MATRIX FOR DIFFERENT CONFIGURATIONS

All of the situations of interest involve transport in one dimension. This requires that we solve the full two dimensional equation of motion of the density matrix. Recall that the second dimension is related to the momentum component in the Wigner formulation. In this section we show results only for the **diagonal** element. The full two dimensional solutions are discussed below.

We point out that if we were interested **only** in time independent solutions it would be possible to obtain them from time independent solutions to Schrodinger's equation. The goals of most of these initial calculations was :

(1) to test the numerical algorithms,

(2) to explore the time dependent behavior to steady state, and

(3) to explore the nature of the boundary conditions.

TIME INDEPENDENT FREE CARRIER SOLUTION

For the free carrier the wavefunction appearing in the density matrix, assuming that we are dealing with equation (9b) is $\psi_k \sim \exp[i(kx - \omega t)]$. Along the diagonal the free particle density is constant and the density matrix in the coordinate representation is constant. As shown later, the off-diagonal elements are dependent upon x and x'.

CARRIERS AND A POTENTIAL BARRIER

The standard textbook problem of a quantum mechanical potential barrier is that of a particle of momentum k incident upon the barrier. The resulting textbook problem proceeds to examine the reflection coefficient and the transmission coefficient and the resulting probability of transmission through the barrier.

The zero bias transient problem we study here is as follows: At time $t = 0$, a uniform density of electrons is created in a structure containing a symmetrically placed barrier. The question we ask is : How do the carriers approach equilibrium?

After the initial transient the problem becomes that of a **stream** of particles incident upon a barrier. If we were interested only in the steady state solution to the problem then we need not solve the equation of motion of the density matrix. Schrodinger's equation would be sufficient. The time dependent multiparticle problem requires the solution to the density matrix. For the transient solution our attention was focused on the dynamics of tunnelling and the role of the device boundaries on the dynamics of the carriers. In

addition some of the cases chosen were directed toward testing the robustness of the numerical algorithm.

The problem here corresponded to tracking the time evolution of a uniform distribution of carriers established in a region that included a potential barrier. The carriers could be placed in the device optically. However, no features of the carrier generation was considered, nor were holes introduced into the discussion. Thus at time t = 0, the system consisted of a uniform distribution of carriers plus a potential barrier. The subsequent physics calls for the carriers to tunnel out of the barrier and propagate toward the boundaries of the device. The structure chosen for this study was a 3000Å long GaAs element with a symmetrically placed potential barrier of 50mev height, and with a width of 200Å. After approximately 15fs, much of the excess carriers in the barrier tunnel out of the barrier and begin to accumulate in the vicinity of the barrier. These carriers propagate toward the physical boundaries of the device, where they are subsequently reflected, and propagate back toward the barrier. In this study, however, because the physical boundaries were far removed from the potential barrier, this was not an issue for much of the calculation.

Reflection off the device boundary is a consequence of the imposed boundary conditions. We have currently implemented a set of boundary conditions that will allow for total transmission through the device boundaries. Any realistic situation would involve different percentages of transmitting and reflection physical boundaries.

The distribution of density in the well 150fs after the start of the calculation is displayed in figure (2), along with a sketch of the barrier structure. We call attention to the negligible space charge within the barrier and the accumulation of carriers on either side of the barrier.

CARRIERS AND A POTENTIAL WELL

This situation is also a classical text textbook problem, in which a particle of momentum k is incident upon a potential well. Here either bound states or scattering states are considered, and the probability of transmission calculated. For the problem studied, we dealt with a situation that is similar to that of the potential barrier, namely:

The zero bias transient problem we study here is as follows: At time t = 0, a uniform density of electrons is created in a structure containing a symmetrically placed well. The question we ask is : How do the carriers approach equilibrium?

As in the potential barrier a uniform distribution of carriers was placed in a structure that consisted of 200Å well, 50mev deep, centrally placed in a structure that was 3000 Å thick. As expected there was a considerable accumulation of carriers in the well, with the massive rearrangement occurring within the first 50 fs after the start of the calculation. What was also observed was that the carrier density in the well was always time dependent. And these calculations were carrier out for up to 750 fs. The nature of these oscillations was examined. Here we note that standard quantum mechanics teaches that for a well with a single descrete energy level, that the carrier density will be highest in the center of the well. For one with two levels there will be lobes that peak near the boundaries of the well. Successively higher levels will result in a different distribution of peaks. For a finite

number of levels, the time dependent interaction of these levels is represented as waves bouncing off these boundaries. Our preliminary analysis indicates that we are observing internal time dependent reflection. Indeed further evidence for these internal reflection was obtained by repeating these calculations with a shallower well of depth 30 mev. The effects of the sidelobes was reduced. **Our calculations appear to be the first to expose numerically the wave nature of the charge distribution in potential wells.** We point out that as a consequence of this time dependence it is SRA's opinion that it should be possible to pump a potential well, or possibly a double barrier diode, with a signal about the resonant frequency and observe a significant frequency dependent variation in the output.

The distribution of charge in the well at 30 f⁻ after the start of the calculation is displayed in figure 3, as is a sketch of the well. Note the accumulation of carriers in the center, and the appearance of some local accumulation of charge at the boundaries of the well. The subsequent time dependence of this result will be explored more fully in the later sections.

CARRIERS AND A DOUBLE BARRIER TUNNELING POTENTIAL

This particular structure was the most complicated of all. It consisted of a 650 Å structure, which included two 230 mev potential barriers each 50 Å thick, separated by 50 Å. While this is a standard AlGaAs/GaAs configuration, and is one that was studied at SRA using the moments of the Wigner function, this structure introduces significant proximity issues, particularly as the physical boundary was only 200 Å away from the tunnelling structure.

The initial condition for this structure was as for the above two problems, i.e., the zero bias transient problem we studied is as follows: At time $t = 0$, a uniform density of electrons is created in a structure containing a symmetrically placed double barrier. The question we ask is : How do the carriers approach equilibrium?

Initially there was a uniform distribution of carriers in the device. The subsequent time development displays the tunnelling of these carriers out of the barriers toward both the physical boundaries and toward the region interior to the barriers. Thus it appears that a steady solution would show the presence of a local accumulation of carriers between the barriers in the absence of any bias. This result would be consistent with that obtained by workers at SRA using the moments of the Wigner distributions function [19]. The situation when the tunnel barrier is in close proximity to the boundary shows that the possibility of reflection off the physical boundary has a significant effect on the distribution of charge within the well. For example, it is found that as the charge in the well is decreasing, there is an accompanying charge packet that travels to the physical boundary. When this wave is reflected off the physical boundary and propagates toward the double barrier structure, the charge distribution in the well starts to increase. This repeats periodically.

We have begun to alter the boundary condition, to allow all of the waves that strike the physical boundary to be transmitted across the boundary. **This is a form of dissipation in that the solutions are no longer time reversible.** Preliminary results indicate that the oscillation in the well in significantly decreased. While these new studies are preliminary, the indicate that proximity effects and conditions at the contacts will exert a first order effect on the behavior of quantum phase based devices.

Figure 4a displays the distribution of charge in the double barrier structure 10fs after the

start of the calculation. We direct attention to the fact that during this short time duration, the carriers are removed from the double barrier and accumulate between the barrier and outside of the barriers. While we note that the density of carriers in the well is below that of the density outside of the well, the situation is reversed at later times. For these boundary conditions, when the excess carriers reaches the boundary, these are absorbed, rather than reflected by the boundary. The situation at 50 fs is displayed in figure 4b.

We also point out that the standard approach to examining transport in resonant tunneling structures, namely the Tsu-Esaki approach, is to ignore the time dependent influence of the contact boundary, as well as the time dependence of the wavefunctions within the interior of the double barrier diode. The result is that the predicted terminal characteristics are likely to be incorrect. **It is therefore not surprising that the agreement between theory and experiment is as poor as it currently is!** We now turn to the more general argument.

EXAMPLES OF THE DENSITY MATRIX AND THE WIGNER FUNCTION

We now consider several examples of the density matrix. The first example is that of a time independent free carrier.

TIME INDEPENDENT FREE CARRIER

For this case the density matrix in the 'state' representation is a classical Boltzmann distribution:

(14) $\qquad \rho_{m,n} = \rho_m \delta_{m,n} = [\exp(-E_m/k_b T)] \delta_{m,n} / \Sigma_\ell \exp(-E_\ell/k_b T)$

For eigenstates representing momentum, $E = \hbar^2 k^2 / 2m$. In one dimension the wavefunction is a one dimensional plane wave solution. The density matrix in the coordinate representation is obtained by replacing the summation by an integral:

(15a) $\qquad \rho(x,x') = (2\sqrt{\pi})(\lambda/L) \int dk \exp[(-\hbar^2 k^2/2mk_b T) + ik(x-x')]$
$\qquad\qquad = (1/L) \exp[-mk_b T(x-x')^2/2\hbar^2]$

or:

(15b) $\qquad \boxed{\rho(x,x') = (1/L)\exp[-\zeta^2/\lambda^2]}$

where:

(16) $\qquad \zeta = (x-x')/2$

and

(17) $\qquad \lambda^2 = \hbar^2/[2mk_b T]$

is the square of the thermal deBroglie wavelength. The density matrix, as given by equation (15b) is constant along lines of constant ζ; that is along lines that are either parallel to the line $x = x'$ or are on the diagonal. The density matrix varies as a Gaussian in a direction normal to the this line. The free particle density matrix in the coordinate representation is

displayed in figure (5), for GaAs parameters. In this diagram abscissa and ordinates are x and x', respectively. The first point we note is that the density matrix for the free particle is real. The second point is that for distances $\zeta > \lambda$, the density matrix becomes negligibly small. Since the expectation values of the operators involve an integration over all values of x and x' the implication is that only those operators that are confined to within a thermal deBroglie wavelength will have an impact on the ensemble average expectation value. We will see later which these operators are. We note that the thermal deBroglie wavelength increases as the temperature decreases. In the figures we suppress the dependence of $\rho(x,x')$ on the distance L.

The Wigner function for the free particle solution is the classical Boltzmann distribution:

(18) $f_W(x,p) = \{(1/[L(2\pi m k_b T)^{\frac{1}{2}}])\} \exp[-p^2/(2m k_b T)]$

The second example is that of carriers confined to a box of length L.

CARRIERS CONFINED TO A BOX OF LENGTH L

For this case the wave functions are subject to the conditions $\psi(x=0) = \psi(x=L) = 0$. With box normalization such that $<x|x> = 1$, the solutions are:

(19a) $\psi(x) = \sqrt{(2/L)} \sin(n\pi x/L)$
$E_n = n^2 \hbar^2 \pi^2/(2mL^2)$

and:

(19b) $\boxed{\rho(x,x') = (2/L) \Sigma [\exp - E_n/kT] \{\sin[n\pi x/L]\sin[n\pi x'/L]\}}$

We do not have a closed form solution to this equation, but we note that this solution is also real. We also note that it is not possible to replace the sum by an integral as we did for the free particle case, we will lose the constraint at $x = x' = L$.

The third case we consider is that of free carrier subject to a constant force .

FREE CARRIER SUBJECT TO A CONSTANT FORCE .

Classical physics teaches that for such a case the density is unchanged, and the time derivative of the momentum is constant. In this case with V(x,t) = -Fx, the solution is (after Iafrate [20]):

(20) $\boxed{\rho(x,x',t) = \rho_0(x,x') \exp[i2F\zeta t/\hbar]}$

where $\rho_0(x,x')$ designates the free particle solution in the absence of any applied potential. The key issue associated with the above solution is that it contains imaginary as well as real components in the solution. The time evolution of the density matrix is shown in figure 6. In displaying the results we recognize that the density matrix for the free carrier is

dependent only on the variable ζ. Thus to examine the time evolution of the free carrier density matrix in the coordinate representation we need only examine the projection of the density matrix onto the principle off-diagonal element.

We see that the entire solution oscillates in time with a frequency determined by the applied force and the spatial variable ζ, except on the diagonal where the density is constant. The oscillation frequency is greatest at the edge of the computational domain. The above implies that **'correlations' between points x and x' as represented by the density matrix, varies in time.** Indeed, examining figure 6, we see that the density matrix, which was originally Gaussian about the diagonal, approaches, in time, a delta function about the origin. The solution will periodically oscillate between a Gaussian and a delta function. Further, we see that the imaginary part of the density matrix is subject to inversion symmetry about the origin. (Any apparent absence of inversion symmetry is 'artist' dependent.)

The Wigner function for this case is:

$$(21) \qquad f_w(x,p)=[1/[L(2\pi mk_bT)^{\frac{1}{2}}]\exp[-(p-Ft/\hbar)^2/(2mk_bT)]$$

This result states that the Wigner function, is displaced by an amount, $Ft/\hbar$, and becomes strongly peaked about this point with increasing time. The mean momentum of this system is always, $Ft/\hbar$. Thus, at least in this case, the peaking of the density matrix about the diagonal appears to be related to the peaking of the Wigner function about the momentum $Ft/\hbar$. Whether this result is isolated or is a result of a general nature, awaits further study.

THE CURRENT MATRIX EQUATION
THE ENERGY MATRIX EQUATION
THE CONSERVATION LAWS IN MATRIX FORM

Seeking solutions to the density matrix, without actually solving the equation of motion of the density matrix, is an interesting exercise, and provides some insight into the general character of the solution, but it is of very limited use. To study real physical problems we need to solve the density matrix, and the solutions must provide a transparent means of discussing the physics. The algorithms we use to solve the density matrix are direct and are discussed below.

The discussion in this section is concerned with the means by which we will interpret the numerical results. Some simple illustrations are included. For purposes of **interpretation** we introduce a change variables. We begin by recognizing that the second derivative operators in equation (10) can be expressed as follows:

$$(22) \qquad \partial^2/\partial x^2-\partial^2/x'^2 = [\partial/\partial x-\partial/\partial x'][\partial/\partial x+\partial/\partial' x]$$

Then with the transformation:

$$(23) \qquad \eta=(x+x')/2, \quad \zeta=(x-x')/2$$
$$x=(\eta+\zeta), \quad x'=(\eta-\zeta)$$
$$\partial/\partial\eta=\partial/\partial x+\partial/\partial x'$$
$$\partial/\partial\zeta=\partial/\partial x-\partial/\partial x'$$

we find:

(24) $\partial^2/\partial x^2 - \partial^2/\partial x'^2 = [\partial/\partial\zeta][\partial/\partial\eta]$

and the equation of motion of the density matrix becomes:

(25) $\partial\rho(\eta+\zeta,\eta-\zeta,t)/\partial t$
            $= (i\hbar/2m)[\partial^2/\partial\eta\partial\zeta]\rho(\eta+\zeta,\eta-\zeta,t)$
            $- (i/\hbar)[V(\eta+\zeta,t)-V(\eta-\zeta,t)]\rho(\eta+\zeta,\eta-\zeta,t)$

We now introduce a number of new terms. But we first recognize that the diagonal component of the density matrix is the density:

(26) $\rho(x) = \rho(x,x)$

We will need to define a current and energy matrix in the coordinate representation. These matricies arise from the expression for the momentum and energy operators in the coordinate representation [21]. In terms of the operators we define a current density matrix in the coordinate representation, j(x,x').

The **elements of the current density matrix** are:

(27a) $j(x,x') = -(\hbar/2mi)[\partial\rho(x,x')/\partial x' - \partial\rho(x,x')/\partial x]$

or:

(27b) $\boxed{j(x,x') = (\hbar/2mi)\partial\rho(\eta+\zeta,\eta-\zeta)/\partial\zeta}$

where the time dependence is suppressed. Note that since $\rho(x,x') = \text{Re}\rho(x,x') + i\text{Im}\rho(x,x')$, and since $j(x,x') = \text{Re}j(x,x') + i\text{Im}j(x,x')$, $\text{Re}j(x,x')$ is obtained from $\text{Im}\rho(x,x')$ and $\text{Im}j(x,x')$ is obtained from $\text{Re}\rho(x,x')$.

For the free particle in the absence of any applied field the current density matrix is **zero** everywhere.

For the **free particle subject to an applied field** the current density matrix is

(28) $j(x,x') = [(\hbar/2mi)\partial\rho_0(\eta+\zeta,\eta-\zeta)/\partial\zeta] + \rho_0 Ft/m$

Along the diagonal $j(x,x) = \rho(x)Ft/m$, which is the classical result.

We now introduce energy density matrix [21]. The **elements of the energy density matrix** are:

(29a) $E(x,x') = -(\hbar^2/8m)[\partial^2\rho/\partial x^2 - 2\partial^2\rho/\partial x\partial x' + \partial^2\rho/\partial x'^2]$

or:

29b)
$$E(x,x')=-(\hbar^2/8m)\partial^2\rho/\partial\zeta^2$$

Here unlike the current density matrix, the real part of the energy density matrix, $ReE(x,x')$ is expressed in terms of $Re\rho(x,x')$ and the imaginary part of the energy density matrix, $ImE(x,x')$, is expressed in terms of $Im\rho(x,x')$.

It is important to note that it is, in principle, possible to solve the equation of motion of the density matrix without any information concerning the current and energy matricies. This would not violate any mathematical precepts. It will, however, be necessary in the course of examining the physics of the problem to constrain the boundary conditions such that specific physical properties related to current and energy are included. This will be apparent in the discussion below.

It is a direct matter to compute the energy matrix **for the free particle in the absence of any applied field.** It is:

(30) $\quad E(x,x')=[(k_bT/2)+2m(\zeta k_bT/\hbar)^2]\rho(x,x')$

Along the diagonal:

(31) $\quad E(x,x)=(k_bT/2)\rho(x)$

which is what is expected on the basis of equipartition. For **the time dependent case of a free particle in an applied field:**

(32) $\quad E(x,x')=-(\hbar^2/2m)\exp[i2F\zeta t/\hbar]$
$\qquad\qquad\times\quad [\partial^2\rho_o/\partial\zeta^2+(4Ft/\hbar)\partial\rho_o/\partial\zeta+(4Ft/\hbar)^2\rho_o]$

and along the diagonal:

(33) $\quad E(x,x)=[(k_bT/2)+(\hbar Ft)^2/2m]\rho(x)$

The energy increases with the square of time as in classical physics.

Finally, the energy matrix for the particle in the box is:

(35) $\quad E(x,x')=(2/L)\Sigma[\exp-E_n/kT]E_n$
$\qquad\qquad\times\{\sin[n\pi x/L]\sin[n\pi x'/L]+\cos[n\pi x/L]\cos[n\pi x'/L]\}$

and along the diagonal:

(36) $\quad E(x,x)=(2/L)\Sigma[\exp-E_n/kT]E_n$

Unlike the results for the free particle in the absence and presence of an applied field the energy matrix for the particle in a box, has no classical counterpart. We now introduce conservation laws.

The matrix form of the equation of continuity: In term of the elements of the current

matrix, the equation of motion of the density matrix can be **rewritten** as follows:

(37) $\partial\rho(\eta+\zeta,\eta-\zeta)/\partial t+\partial j(\eta+\zeta,\eta-\zeta)/\partial\eta$
         $=(1/i\hbar)[V(\eta+\zeta)-V(\eta-\zeta)]\rho(\eta+\zeta,\eta-\zeta)$

Equation (37) teaches that the density matrix equation conserves particles. Further, along the diagonal, $\zeta=0$, the above differential equation is real and yields the well known continuity equation:

(38) $\partial\rho(x)/\partial t+\partial j(x)/\partial x=0.$

We next consider **the matrix form of the force balance equation**. To do this we take derivatives of the density matrix equation (25) with respect to ζ. We find, with $j(x,x')=p(x,x')/m$:

(39) $\partial p(x,x')/\partial t+2\partial E(x,x')/\partial\eta=$
         $-(1/2)\rho(x,x')\partial[V(x)-V(x')]/\partial\zeta-(i/\hbar)[V(x)-V(x')]p(x,x')$

Equation (37) is a generalized matrix form of Newton's Law. We point out, however, that if the density matrix is real, the real part of the above equations reduces to:

(40) $2\partial E(x,x')/\partial\eta+(1/2)\rho(x,x')\partial[V(x)-V(x')]/\partial\zeta=0,$

which as we will see is consistent, along the diagonal and under special restrictions, with the time independent Schrodinger equation. Before doing this we note that:

(41) $\partial/\partial\zeta[V(\eta+\zeta)-V(\eta-\zeta)]=[V'(\eta+\zeta)+V'(\eta-\zeta)]$

where the prime denotes differentiation with respect to the argument of the respective functions. Then without confining the discussion to the real part of equation (39), we find that along the diagonal we are able to retrieve an **ostensibly** classical result:

(42) $\partial p(x,x)/\partial t+2\partial E(x,x)/\partial x=-\rho(x,x)\partial V(x)/\partial x$

Since p(x,x') and E(x,x') are density dependent quantities, equation (42) is Newton's Law for particle density.

Is there any quantum mechanics in equations (38) and (42)? For the case of the free particle all of the quantum mechanical treatment yields only the classical result. Thus **in the absence of any quantum mechanical constraints the above analysis is the mathematical equivalent of opening a peanut shell with a sledge hammer.** But there is quantum mechanics, and we demonstrate this by example. Consider the case where the density matrix is **real**. In addition assume for **simplicity** that we can break up the density matrix as $\rho(x,x')=\sqrt{\rho(x)}\sqrt{\rho(x')}$, a situation that is always true for a pure state. Then along the diagonal:

(43) $E(x,x)=-(\hbar^2/8m)2\{\sqrt{\rho}\partial^2\sqrt{\rho}/\partial x^2-(\partial\sqrt{\rho}/\partial x)^2\}$

Introducing the **quantum potential** [22], discussed above:

(44) $Q = -[\hbar^2/2m][1/\sqrt{\rho}]\partial^2\sqrt{\rho}/\partial x^2$

it is a direct matter to show that $\partial E(x)/\partial x = (\rho/2)\partial Q/\partial x$. In this case, equation (41)

(45) $\partial(Q+V)/\partial x = 0.$

or:

(46) $-[\hbar^2/2m]\partial^2\sqrt{\rho}/\partial x^2 + V\sqrt{\rho} = \text{Constant} \times \sqrt{\rho}$

which is Schrodinger's equation for a real wavefunction. We are now at the point were we can begin to discuss the development of boundary conditions to the solution of the equation of motion of the density matrix.

BOUNDARY CONDITIONS AND SOLUTIONS

The purpose in illustrating the importance of boundary condition is two-fold:

First, because the initial discussion of Frensley [18] was incorrect, and may have hampered the introduction of the density matrix as a viable means of solving transport problems, it is important to establish that conditions can be applied that permit solutions.

Second, it is necessary to connect the boundary conditions to those of others as well as to physical boundary conditions. We illustrate this only for the simple case of the standard wave equation. The discussion that follows can be found in textbooks.

We consider the simple case of a time independent problem with zero applied potential. **For this case the density matrix equation reduces to the standard wave equation:**

(47) $[\partial^2/\partial x^2 - \partial^2/\partial x'^2]\rho(x,x') = 0$

The following conditions will yield a unique solution:

(48) $\rho(x,0) = \alpha(x)$
$\partial\rho(x,x')/\partial x'\,|_{x'=0} = \beta(x)$
$\rho(0,x') = \mu(x')$
$\rho(L,x') = \nu(x')$

Because each of the above conditions is related to the Fourier transform in momentum space of the Wigner function, the above conditions are equivalent to momentum boundary conditions on the Wigner function. Thus if the Wigner function specified the distribution of momentum on the boundaries, the above conditions would reflect this conditions. Thus, physics is contained in the boundary conditions.

The domain of the solution is:
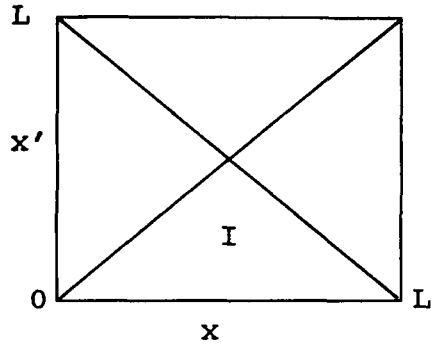
Using the variable changes of equation (23), we can write the general solution of the wave equation as:

(49) $$\rho(x,x')=f(\eta)+g(\zeta)$$

Thus **two** linearly independent solutions are required for the wave equation. Frensley's [18] mistake was that he imposed boundary conditions on only one of the solutions. Thus his general solution was not unique and he was not able to obtain a computationally converged solution.

Returning to the conditions on the differential equation, the boundary and initial conditions can be reexpressed as:

(50) $$\alpha(x)=f(x)+g(x)$$
$$\beta(x)=f'(x)-g'(x)$$
$$\mu(x')=f(x'/2)+g(-x'/2)$$
$$\nu(x')=f(\{L+x'\}/2)+g(\{L-x'\}/2)$$

where the prime on the functions denote derivatives. It is important to recognize that the range of α and β is $0<x<L$. Now it is direct that:

(51) $$\alpha'(x)=[f'(x/2)+g'(x/2)]/2$$

Therefore:

(52) $$f'(x/2)=\tfrac{1}{2}\alpha'(x)+\tfrac{1}{2}\beta(x)$$
$$g'(x/2)=\tfrac{1}{2}\alpha'(x)-\tfrac{1}{2}\beta(x)$$

Integrating:

(53) $$f(x/2)=\tfrac{1}{2}\alpha(x)+\tfrac{1}{2}\int_0^x\beta(\xi)\,d\xi$$
$$g(x/2)=\tfrac{1}{2}\alpha(x)-\tfrac{1}{2}\int_0^x\beta(\xi)\,d\xi$$

and

(54) $$\rho(x,x')=\tfrac{1}{2}\alpha(x+x')+\tfrac{1}{2}\alpha(x-x')+\tfrac{1}{2}\int_{x-x'}^{x+x'}\beta(\xi)\,d\xi$$

It is important to recognize that both α and β are defined in the range $0<x<L$. Thus equation (54) defines solutions to the density matrix, but because of this constraint these

solutions are restricted to **the range of values of** ρ **(x,x') to triangle 'I' of the accompanying figure.** Within this triangle we can illustrate some features of the solution, where we note that there is as yet no explicit dependence of the solution on the vertical boundary conditions μ(x') and ν(x'). We also note that along the bounding diagonal, x = x'(or ζ = 0), and off diagonal elements, x + x' = L (or L = 2η), the solutions are, respectively:

(55) ρ (x,x) =½α (2x) +½α (0) +½$\int_0^{2x}\beta$ (ξ) dξ

(56) ρ (x,L−x) =½α (L) +½α (2x−L) +½$\int_{2x-L}^{L}\beta$ (ξ) dξ

We now consider the solution in the interior of the computation domain. We consider the domain as before:
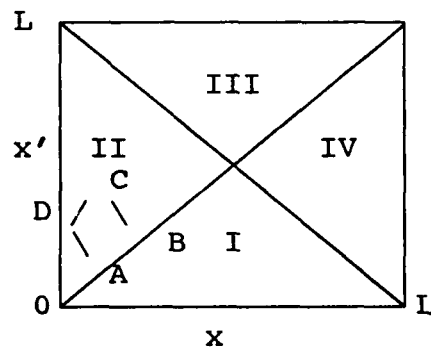
In the above figure, the line segments DC and AB are parallel, as are the line segments AD and BC. These line segments are segments of characteristic lines [ ]. We recognize that the function f(η) is constant along the off-diagonal; while the function g(ζ) is constant along the diagonal. For purposes of illustration:

along the line segment AB, g(ζ) = a;
along the line segment BC, f(η) = b;
along the line segment CD, g(ζ) = c;
along the line segment DA, f(η) = d.

Thus at the points A,B,C,D:

(57) ρ (A) =a+d
           ρ (B) =a+b
           ρ (C) =c+b
           ρ (D) =c+d

and:

(58) ρ (A) +ρ (C) =ρ (B) +ρ (D)

For a point c(x,x') in domain II defined by 0≤x≤½ and x≤x'≤ < -x

$$(59) \qquad A \equiv \left[\frac{x'-x}{2} \, , \, \frac{x'-x}{2} \right]$$

$$B \equiv \left[\frac{x'+x}{2} \, , \, \frac{x'+x}{2} \right]$$

$$C \equiv (x,x')$$

$$D \equiv (0, \ x'-x)$$

$$(60) \qquad \rho(C) = \mu(x'-x) + \frac{1}{2} \left[\alpha \ (x'+x) - \alpha(x'-x) \right] + \frac{1}{2} \int_{x'-x}^{x'+x} \beta(\zeta) \, d\zeta$$

Thus with a function boundary on the left hand side of the domain we can find a complete solution to the density matrix in region II. In a similar manner we can find a solution everywhere in region IV, based upon the conditions on the right hand boundary.

We note that in constructing solutions in region II, these solutions were based upon the solutions on the lines bounding region II. Thus for the solutions in region IV, these will be based upon solutions on the triangular boundaries of the region. For solutions in region III, we can construct solutions from three points providing one of the points is the intersection of the principle diagonal and off-diagonal line segments.

We need not necessarily deal with the above conditions alone, we could choose to replace the conditions of equation (48) with:

$$(61) \qquad \rho(x,0) = \alpha(x)$$
$$\partial \rho(x,x')/\partial x' |_{x'=0} + \partial \rho(x,x')/\partial x |_{x'=0} = \beta(x)$$
$$\rho(0,x') = \mu(x')$$
$$\rho(L,x') = \nu(x')$$

and a unique solution would also be obtained.

The task of the program to which we have been involved requires that we find a solution to the **time dependent** problem that is consistent with the solutions to the time independent problem. We have achieved success in this area. We illustrate this further with an expansion of the results for the solution to the density matrix in the potential well, discussed earlier.

TWO DIMENSIONAL TIME DEPENDENT SOLUTION TO THE
DENSITY MATRIX FOR CARRIERS AND A POTENTIAL WELL

This discussion is an extension of that in connection with figure 3. Here in figure 7 and 8 we present a sequence of projections of the real and imaginary part of the density matrix at successive instants of time. In figure 7, we display the real part of the density matrix as a function of x and x'. At a time t = 15fs, along the diagonal where x = x' we see a uniform distribution of carriers approaching the boundary and a large accumulation of carriers within the well. During the early time phase there is the development of the carrier reflection off the interior of the potential well, figures 7b through 7d. At a later time with larger time increments we begin to see the presence of structure in the solution along the off diagonal in the vicinity of the potential well. In time the wave propagates toward the edge of the domain and reaches it at a time of approximately 500 fs. Further increases in time show the evolution of the charge distribution, but at all times it remains dominated by the increase in charge in the potential well.

TWO DIMENSIONAL TIME DEPENDENT SOLUTION TO THE DENSITY
MATRIX FOR THE DOUBLE BARRIER TUNNELING STRUCTURE

This calculation which is an extension of that discussed in connection with figure 4 was initiated approximately one week before the start of this document. In figure 8 we display the real part of the density matrix at three instants of time. In figure 8a, the presence of the double barrier results in an accumulation of charge between the barriers as well as charge on either side of the barrier. The display along the diagonal, illustrated in figure 4, is the same as that in figure 8. We also note the structure along the off-diagonal. At a time of 30 fs, the propagating wave reaches the boundary, and we see an apparent peak in the charge at the center. In figure 8c we see an apparent decrease in the charge in the well and an increase in the charge at the bounding edges.

Figure 9 displays the distribution of velocity for this structure. At 10fs, figure 9a, the velocity at the boundaries, along the diagonal is zero, and reflects the fact that the propagating charge disturbance has not yet reached the boundary. Note that the velocity is the velocity flux, or current. At time t = 30 fs the disturbance has reached the boundary as indicated by the locally higher values of velocity along the diagonal near the edges of the structure. At 60 fs the velocity is negative along the left hand boundary indicating that the waves have been reflected off the boundary.

The velocity is obtained from the derivative of the imaginary part of the density matrix in the off diagonal direction. The imaginary part of the density matrix is displayed in figure 10. The value of the imaginary part of the density matrix at 10 fs is displayed in figure 10a. We see that it is zero near the edges, indicating again that the disturbance has not yet reached the boundary. At 30 fs, the values of the imaginary part of the density matrix indicate again the propagating disturbance. Figure 10c shows the disturbance at a later time. The important point to note here is that the peaks have altered position indicating that the velocity has changed sign.

In figure 11 we display the real part of the energy matrix. Along the diagonal and in steady state the potential plus E(x,x) is a constant. These calculations have not reached a steady state. The calculations point to the fact that during the transient the energy appears to be

greater where the density is highest, and that at 30 fs when the carriers are near the edge of the boundaries there is a large kinetic energy component.

Finally, we have begun to explore the role of the boundary on the results. In a set of calculations we began to let a certain fraction of the carriers be transmitted through the device boundaries. The result of this calculation at 60fs, display a more isolated distribution of charge in the well. This clearly is an avenue of research.

3. RECOMMENDATIONS FOR FUTURE WORK

The above discussion sets the stage for the type of research that needs to be performed in the future. The density matrix offers an excellent means of studying the physics of quantum phase based devices. It is can be modified:

(1) to treat heterostructures,

(2) include Poisson's equation,

(3) treat transient transport in a variety of different III-V material configurations,

(4) and treat transient transport in silicon and germanium heterostructures.

It is capable of studying multiple barrier diodes (more than two), and of including dissipation. While we are currently treating dissipation at the boundaries, dissipation can be treated, phenomenologically through a relaxation time approximation. Here we would replace the left hand side of the governing partial differential equation (10) with the terms:

$$(62) \qquad \partial \rho(x,x',t)/\partial t + \rho(x,x',t)/\tau$$

Thus it is capable of treating the device physics of one dimensional structures.

To deal with the device physics of two-dimensional structures, such as A-B devices, requires that a $(4 + 1)$ dimensional partial differential equation be solved. This equation is:

$$(63) \qquad \begin{aligned} &\partial \rho(\mathbf{r},\mathbf{r}',t)/\partial t \\ &= (i\hbar/2m)[\partial^2/\partial \mathbf{r}^2 - \partial^2/\partial \mathbf{r}'^2]\rho(\mathbf{r},\mathbf{r}',t) \\ &- (i/\hbar)[V(\mathbf{r},t) - V(\mathbf{r}',t)]\rho(\mathbf{r},\mathbf{r}',t) \end{aligned}$$

While workers at SRA are **confident** in their ability to develop a transient, four dimensional solution algorithm for such a system of equations, it is apparent that such a procedure would be computationally intensive, and several stages would be required for implementing this.

References

[1] L. Esaki, and R. Tsu, IBM Journal of Research and Development, 61, 1970.

[2] L. L. Chang, L. Esaki and R. Tsu, Appl. Phys. Lett. 24, 593 (1974).

[3] T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C. D. Parker and D. D. Peck, Appl. Phys. Lett. 43, 588 (1983).

[4] A. R. Bonnefoi, R. T. Collins, T. C. McGill, R. D. Burnham and F. A. Ponce, Appl. Phys. Lett. 46, 285 (1985).

[5] S. Ray, P. Ruden, V. Sokolov, R. Kolbas, R. Boonstra and J. Williams, Appl. Phys. Lett. 48, 1666 (1986).

[6] C. I. Huang, M. L. Paulus, C. A. Bozada, S. C. Dukley, K. R. Evans, C. E. Stutz, R. L. Jones and M. E. Cheney, Appl. Phys, Lett., 51, 121 (1987).

[7] S. C. Kan, H. Morkoc and A. Yariv, Appl. Phys. Lett. 52, 2250 (1988).

[8] S. Luryi, Appl. Phys. Lett. 47, 70 (1985).

[9] M. O. Vassell, Johnson Lee and H. F. Lockwood, J. Appl. Phys. 54, 5206 (1983).

[10] M. Cahay, M. McLennan, S. Datta and M.S. Lundstrom, Appl. Phys. Lett. 50, 612 (1987).

[11] K. M. S. V. Bandara and D. D. Coon, Appl. Phys. Lett. 53 (19), 1865 (1988).

[12] B. Jogai et al. (preprint 1989).

[13] T. P. E. Broekaert, Wai Lee and Clifton G. Fonstad, Appl. Phys. Lett. 53 (16), 1545-1547 (1988).

[14] S. Datta, M. R. Mellcoh, S. Bandyopadyay and M. S. Lundstrom, Appl. Phys. Letts. 48, 487 (1986).

[15] H. L. Grubin, D. K. Ferry, K. R. Gleason, Solid State Electronics, 23, 157 (1980).

[16] N. C. Kluksdahl, A. M. Kriman and D. K. Ferry, to appear in Phys. Rev. B.

[17] E. P. Wigner, Phys, Rev. 40, 749 (1932).

[18] W. R. Frensley, J. Vac. Sci. Technol, B3, 1261 (1985).

[19] H. L. Grubin and J. P. Kreskovsky, Solid State Electronics, 32, 1071, (1989).

[20] G. J. Iafrate, Private Communication.

[21] H. L. Grubin, To be Published.

[22] C. Philippidis, D. Bohm and R. D. Kaye, IL Nuovo Cimento, 71B, 75

dc

V_D (VOLTS)

PULSED

V_D (VOLTS)
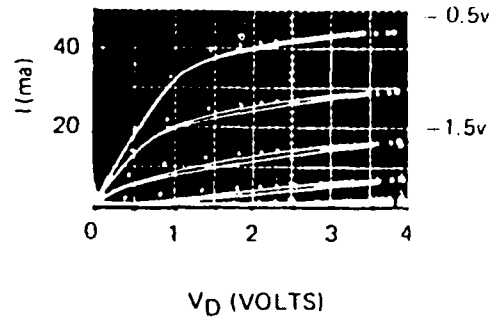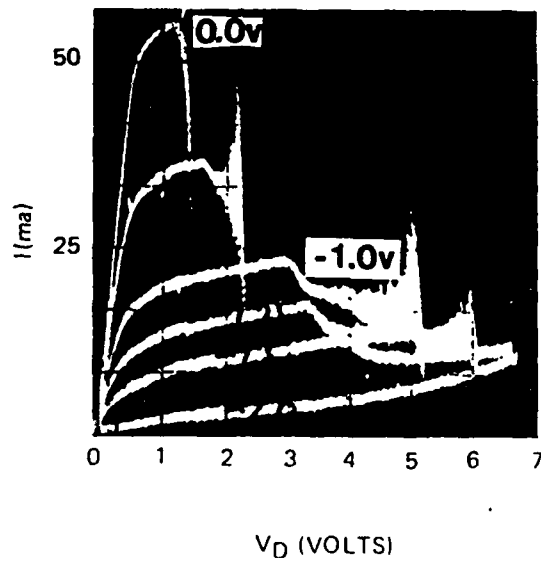
Figure 1.   (a) d.c. and (b) pulsed sampling scope measurements
            of the current voltage characteristics for a gallium
            arsenide FET. [Ref.   ].
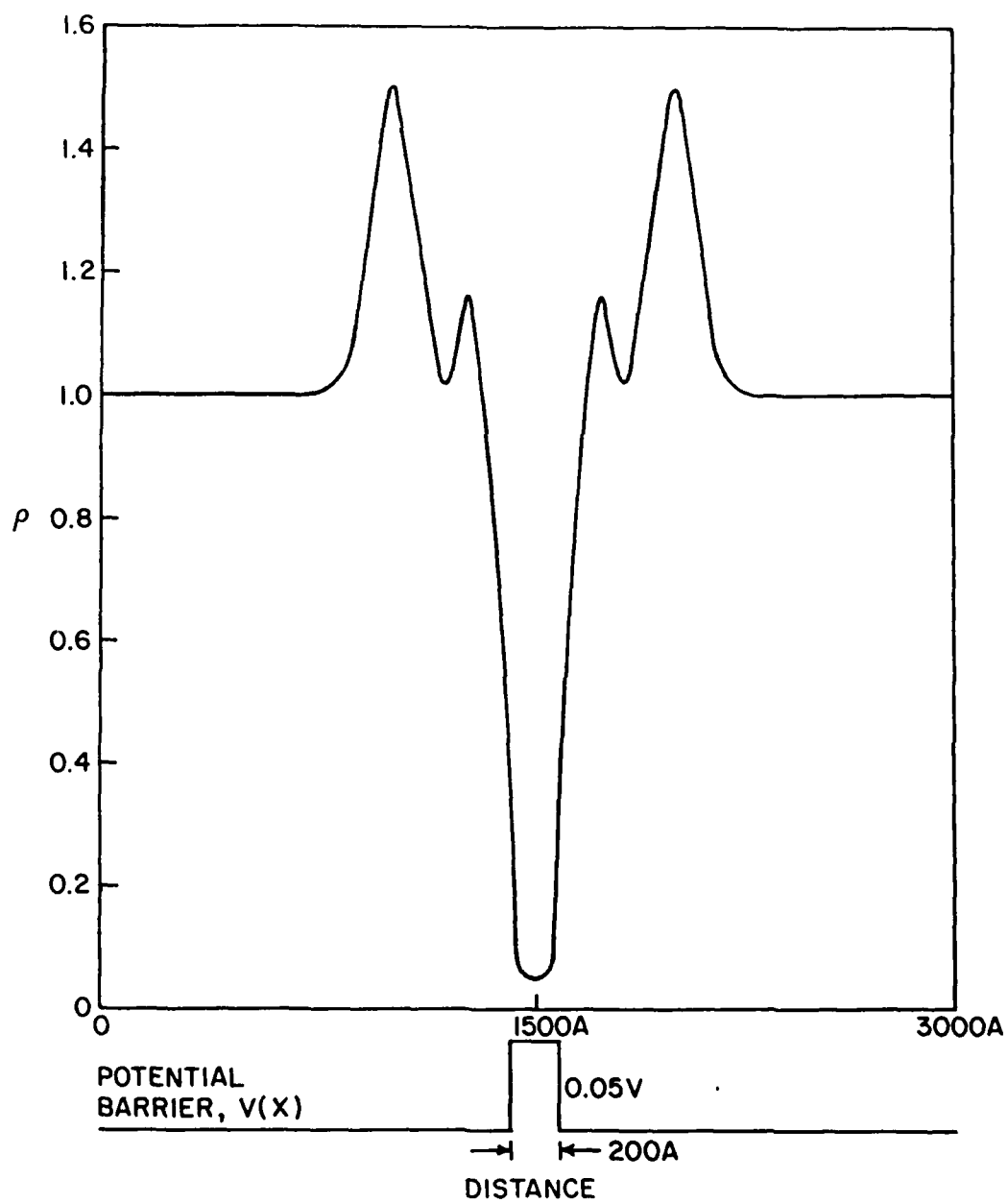
Figure 2. Diagonal component of the density matrix, or the
carrier density, at 150 fs, for a distribution of
carriers interacting with a potential barrier.
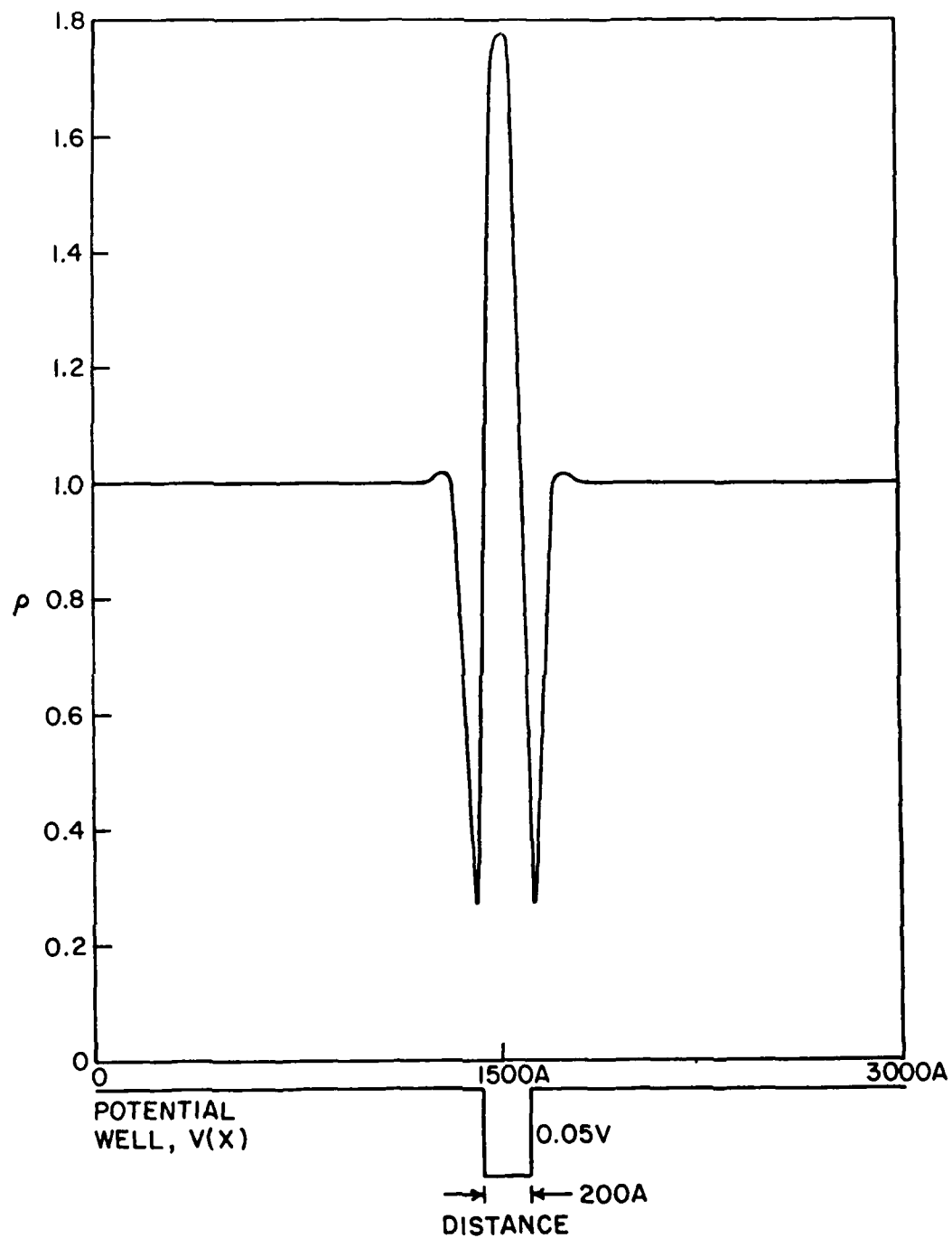
Figure 3.   Diagonal component of the density matrix, or the
            carrier density, at 30 fs, for a distribution of
            carriers interacting with a potential well.
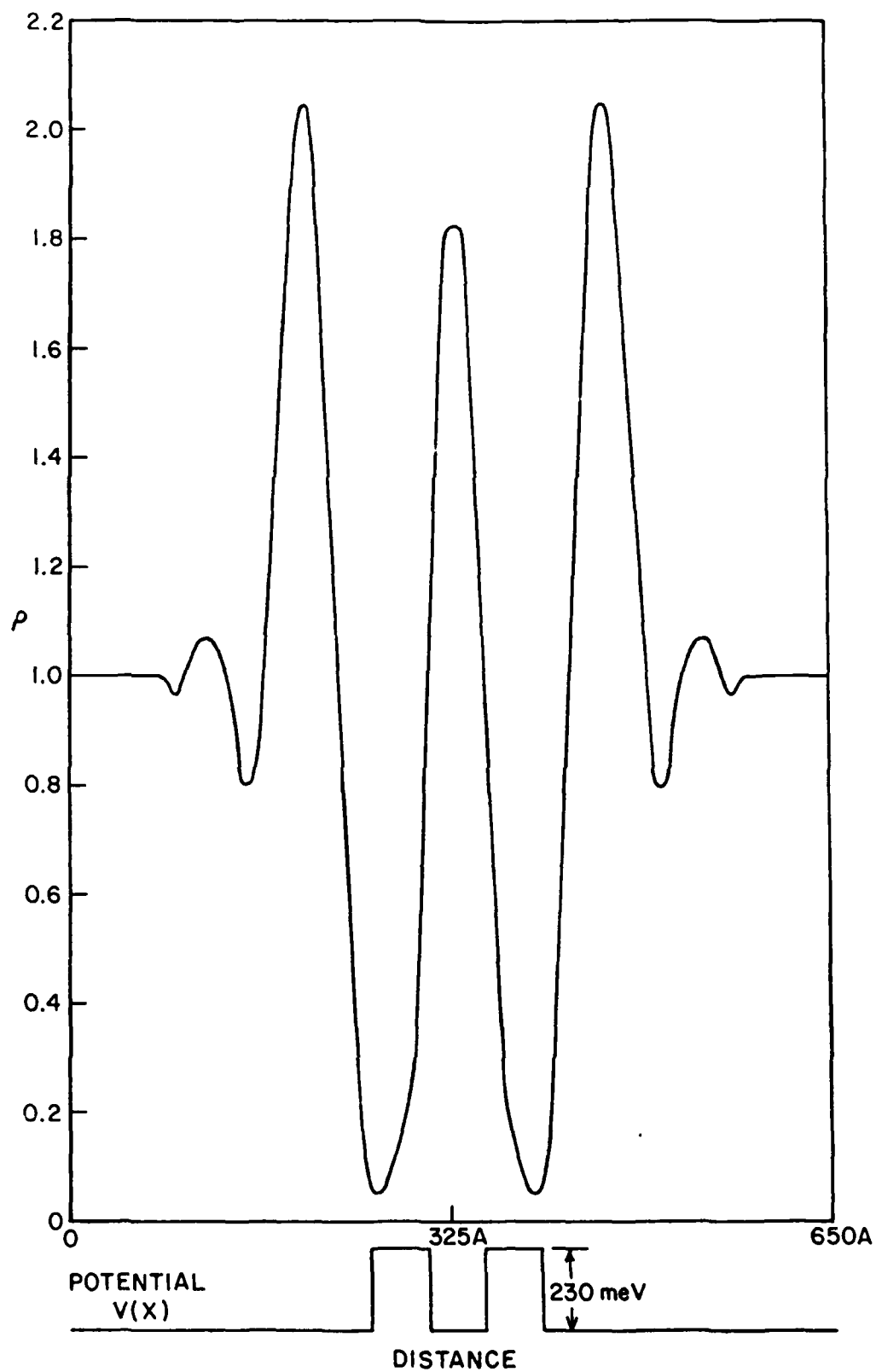
Figure 4a. Diagonal component of the density matrix, or the carrier density, at 10 fs, for a distribution of carriers interacting with a double barrier tunneling structure.
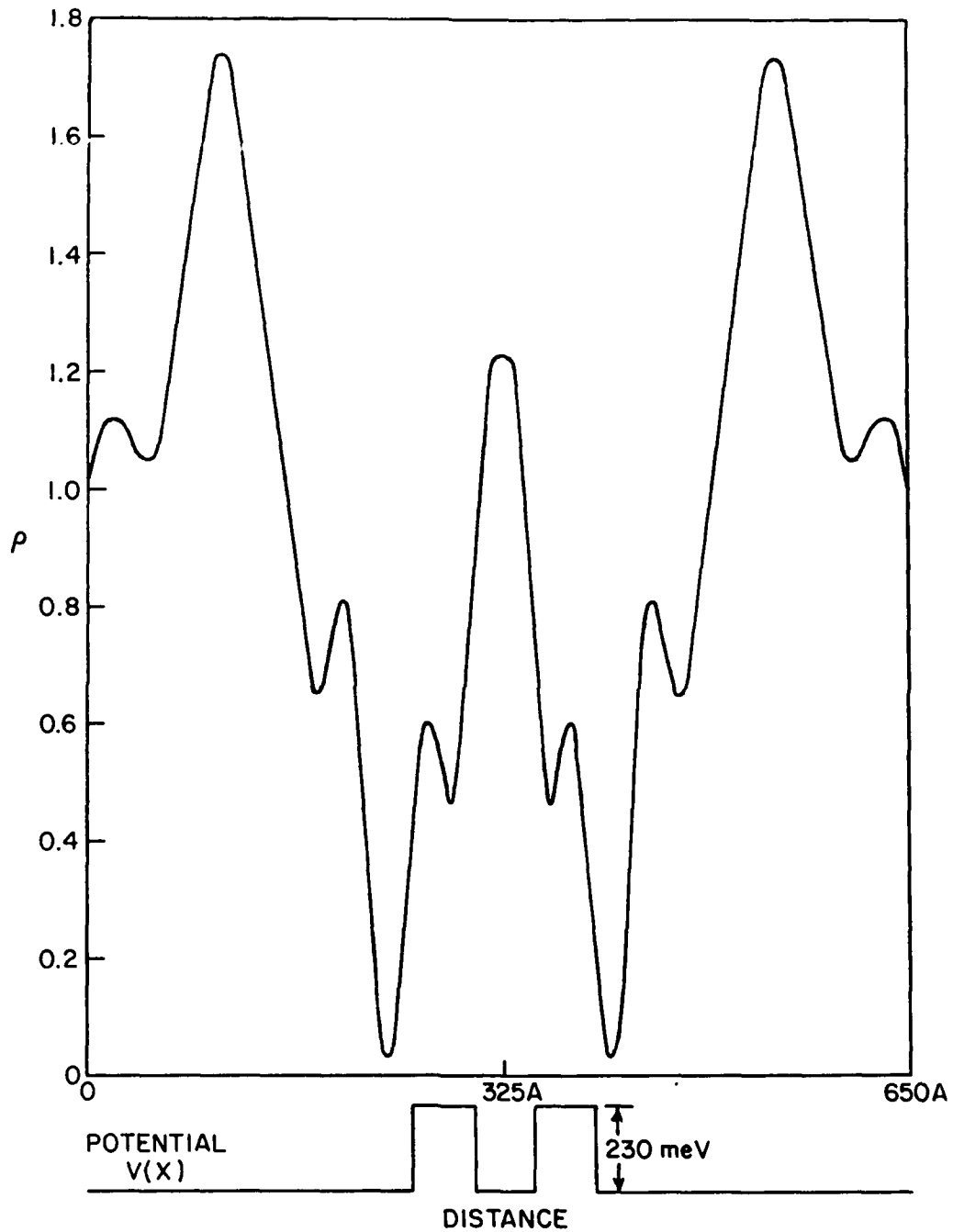
Figure 4b. Diagonal component of the density matrix, or the carrier density, at 50 fs, for a distribution of carriers interacting with a double barrier tunneling structure.
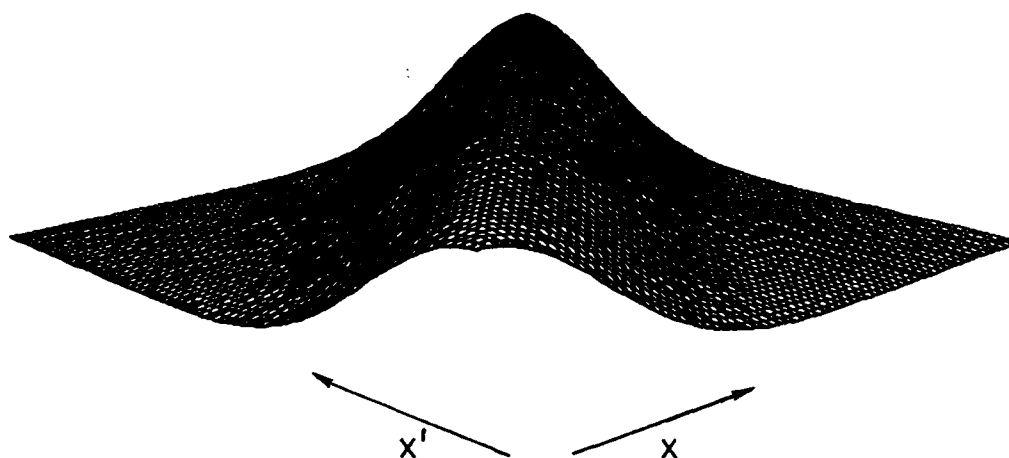
Figure 5a. Free barrier density matrix in the coordinate
            representation corresponding to equation 15b.
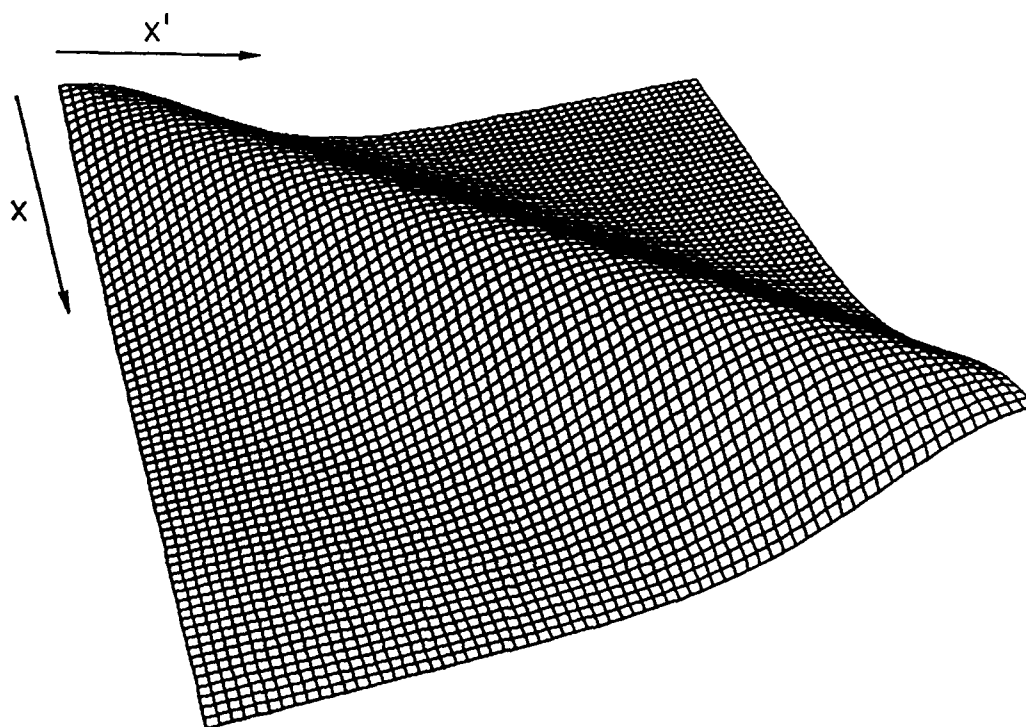
Figure 5b. Carrier density matrix in the coordinate representation corresponding to equation 15b. Figure 5b is a different projection of Figure 5a.

75 fs

300 fs

150 fs

375 fs
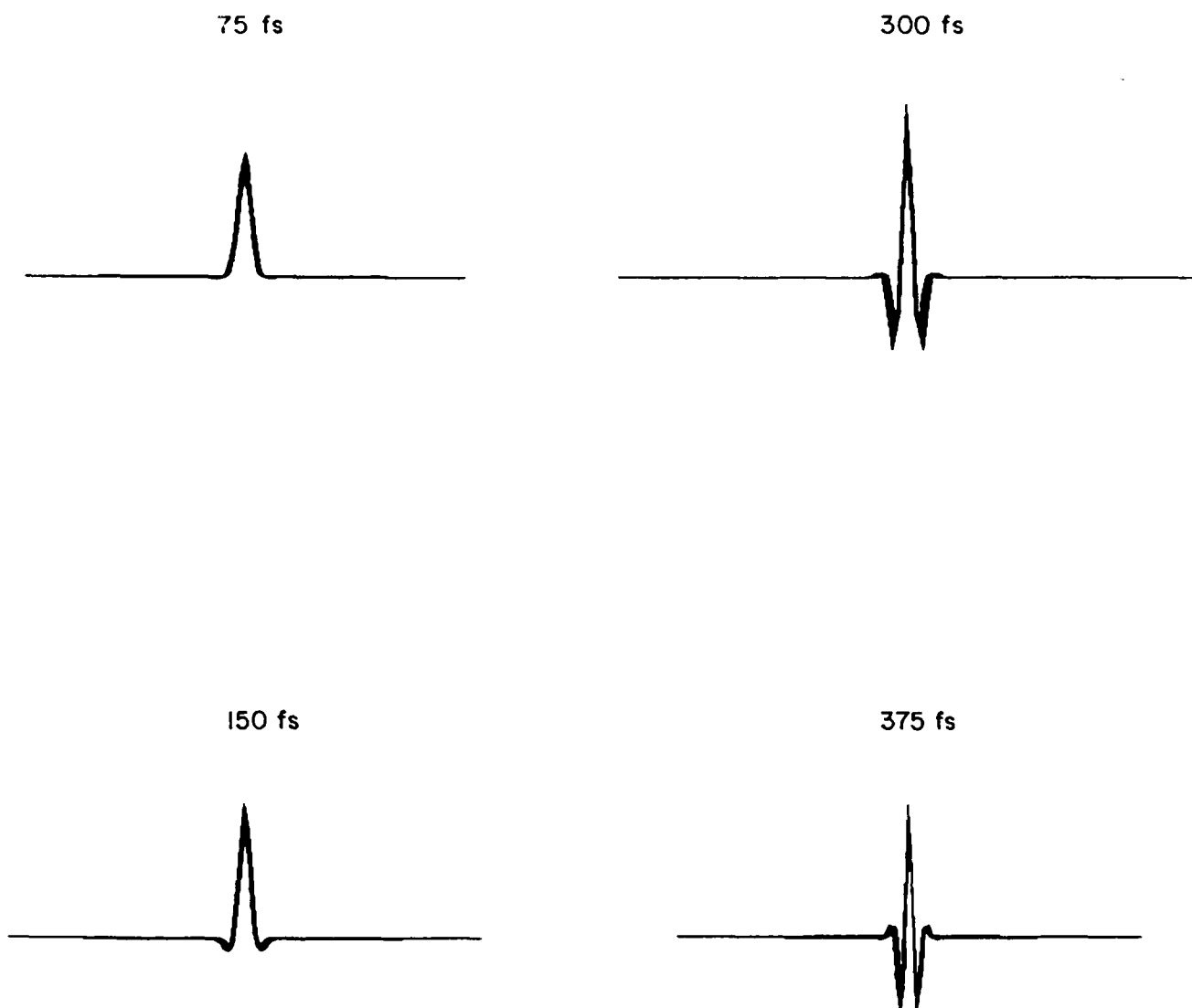
Figure 6a. Projection of $\mathrm{Re}\rho(x,x',t)$ at different times for $V(x,t) = -\mathcal{F}x$.

75 fs

300 fs

150 fs

375 fs
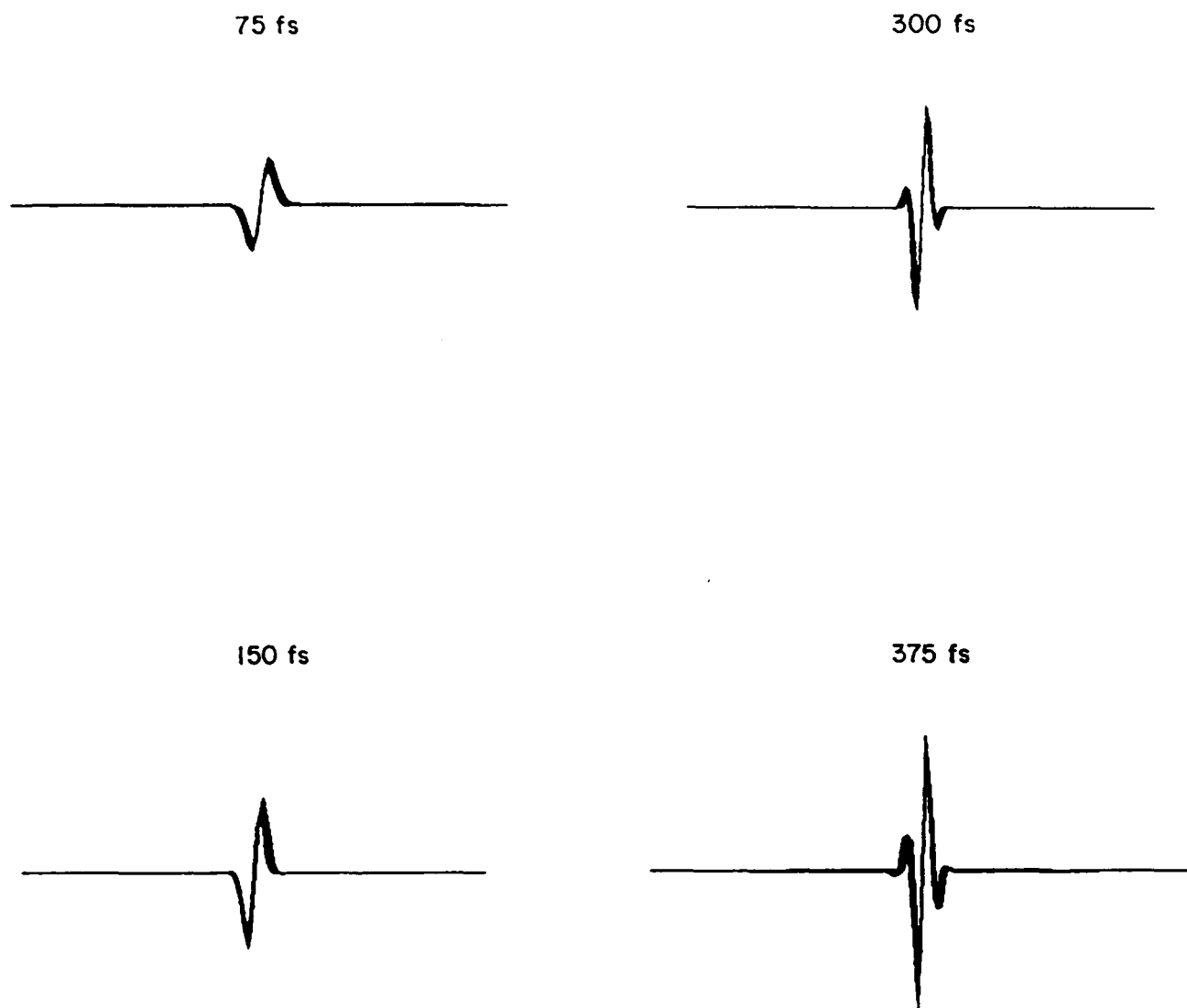
Figure 6b. Projection of Imρ(x,x',t) at different times for V(x,t) = -𝒥x.
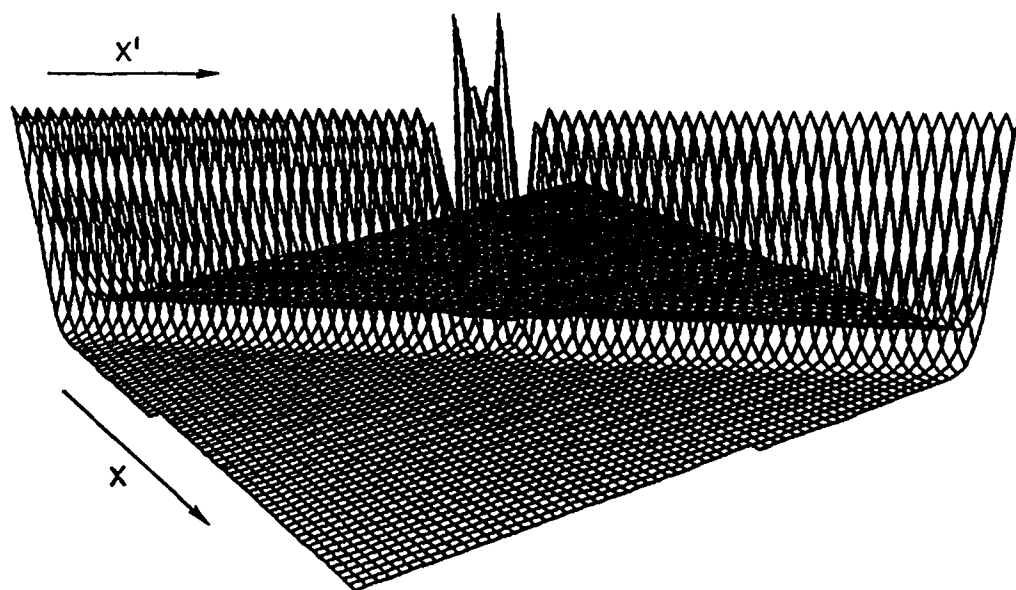
Figure 7a. Reρ(x,x',t) for the potential well of Figure 3. t = 15 fs.
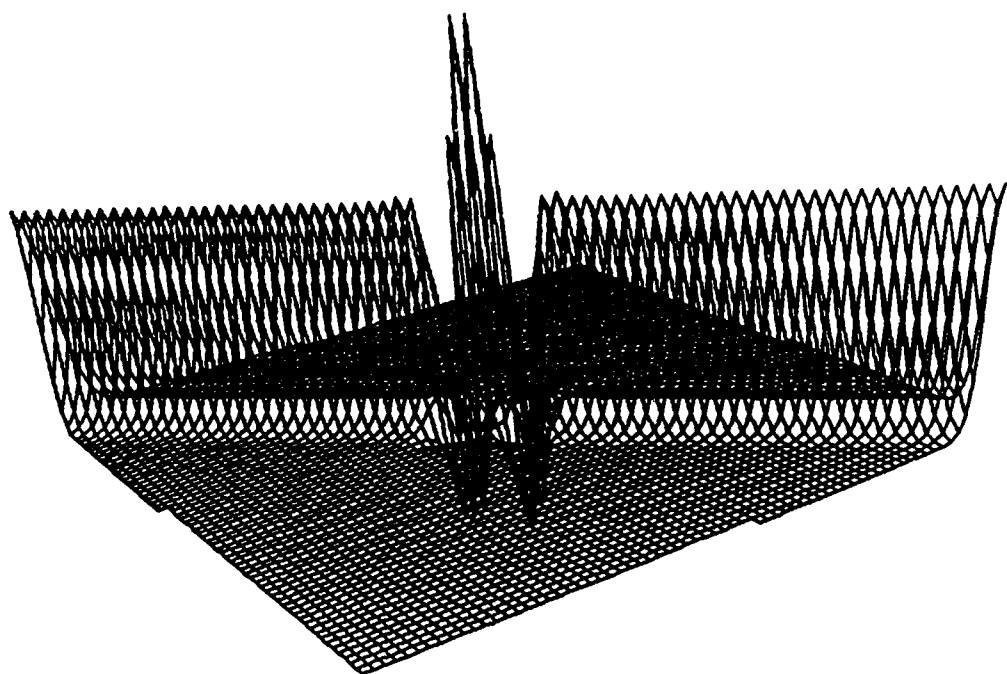
Figure 7b. Reρ(x,x',t = 30 fs).
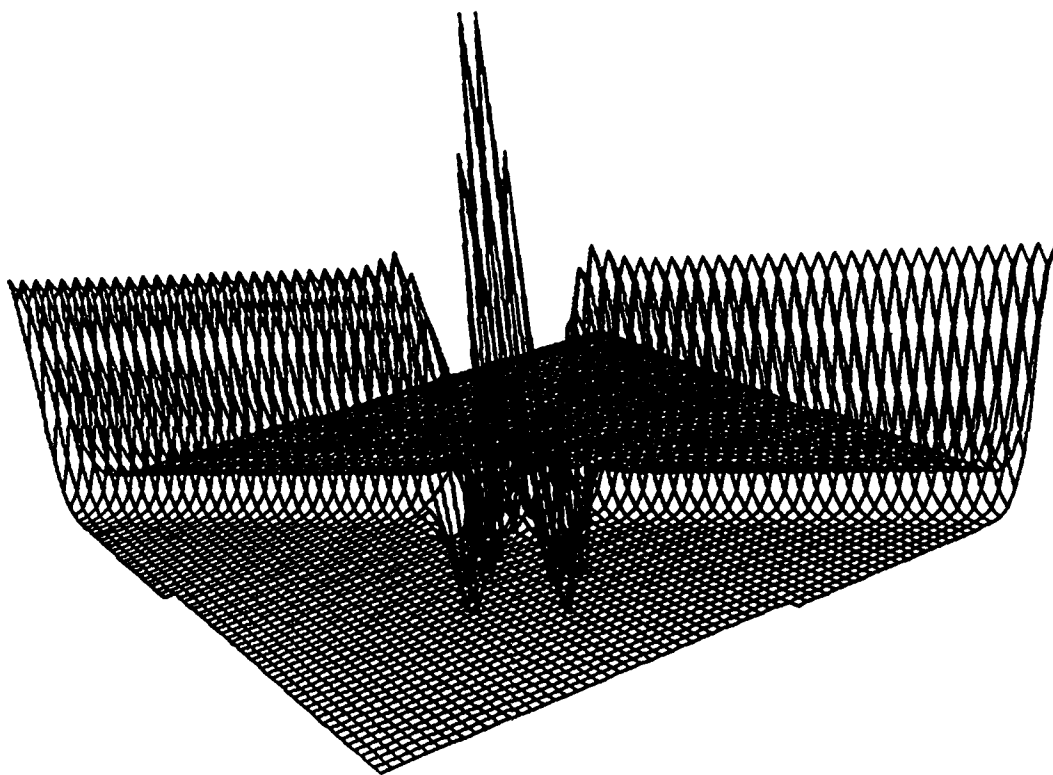
Figure 7c. Reρ(x,x',t = 60 fs).
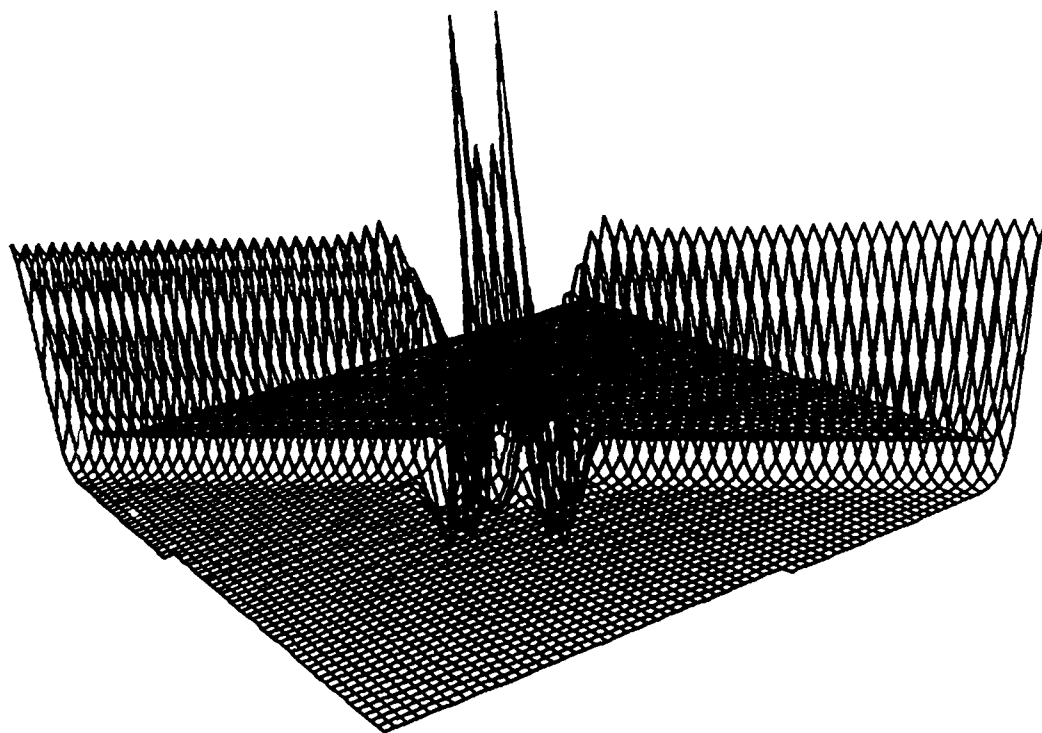
Figure 7d.   Rep$(x,x',t = 75$ fs$)$.
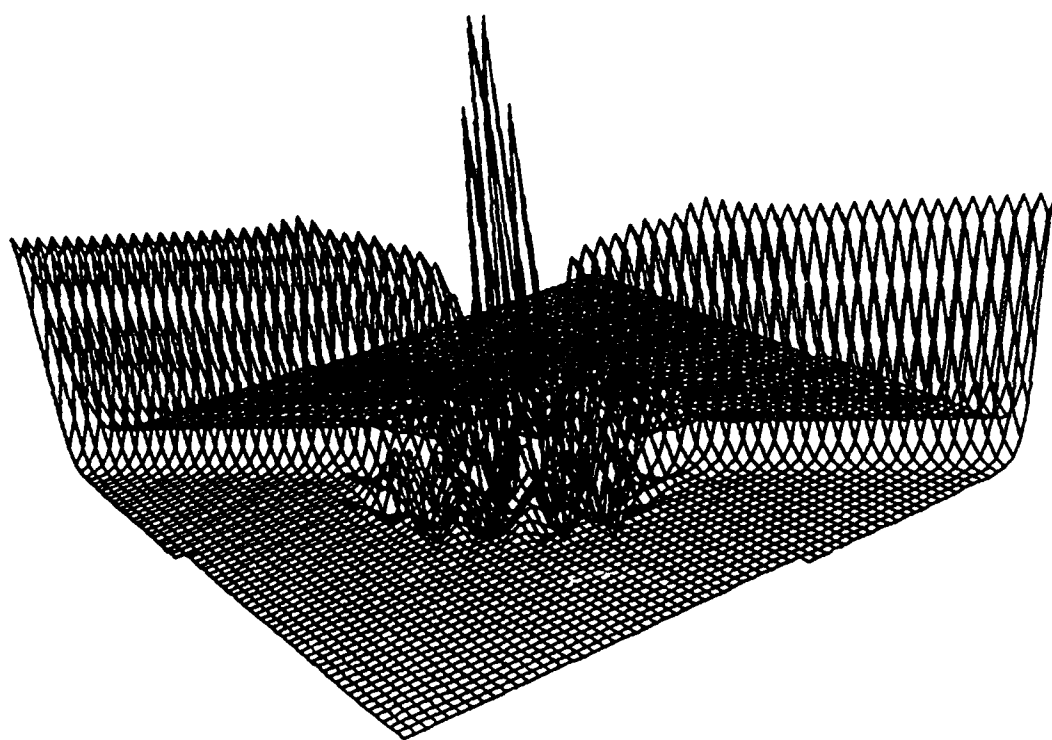
Figure 7e. Re$\rho(x,x',t = 150$ fs).
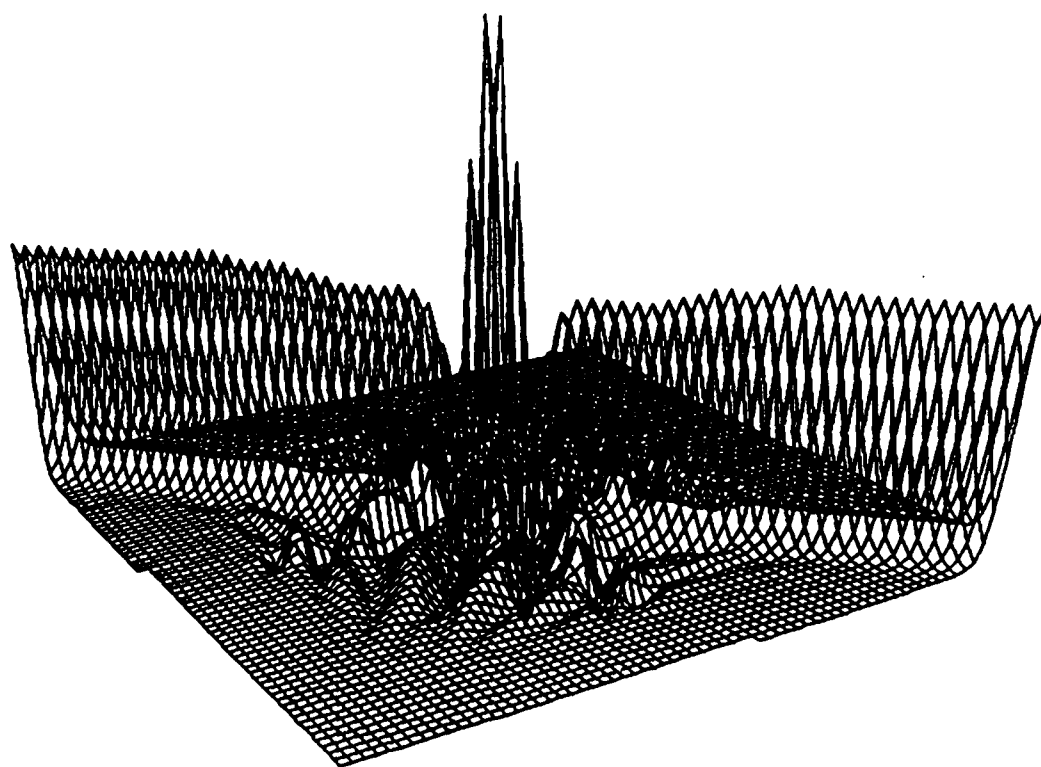
Figure 7f. Reρ(x,x',t = 225 fs).
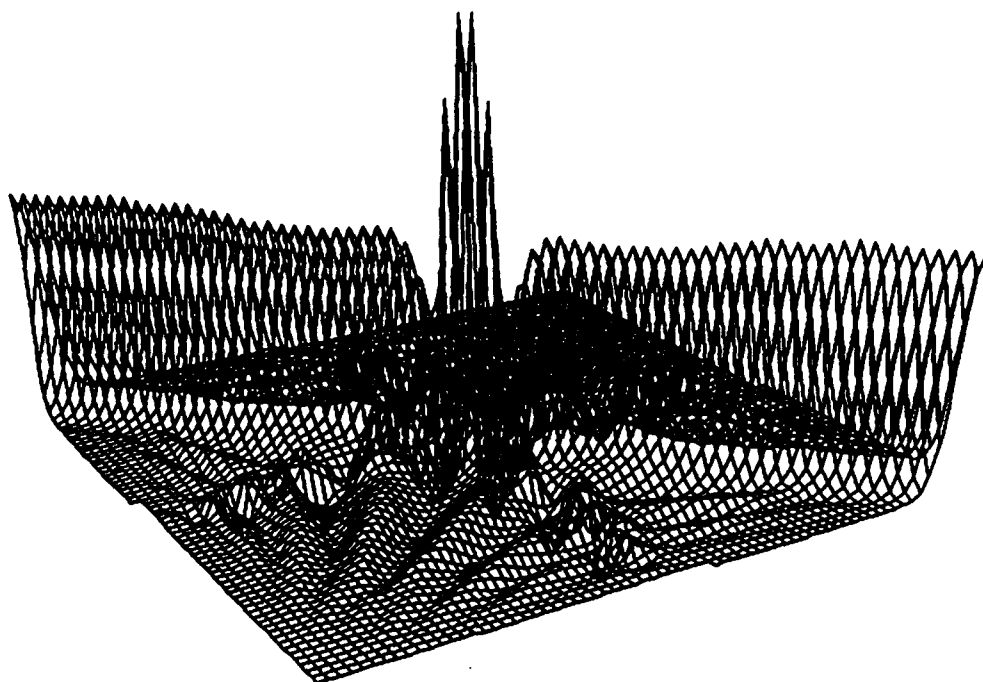
Figure 7g. Reρ(x,x',t = 300 fs).
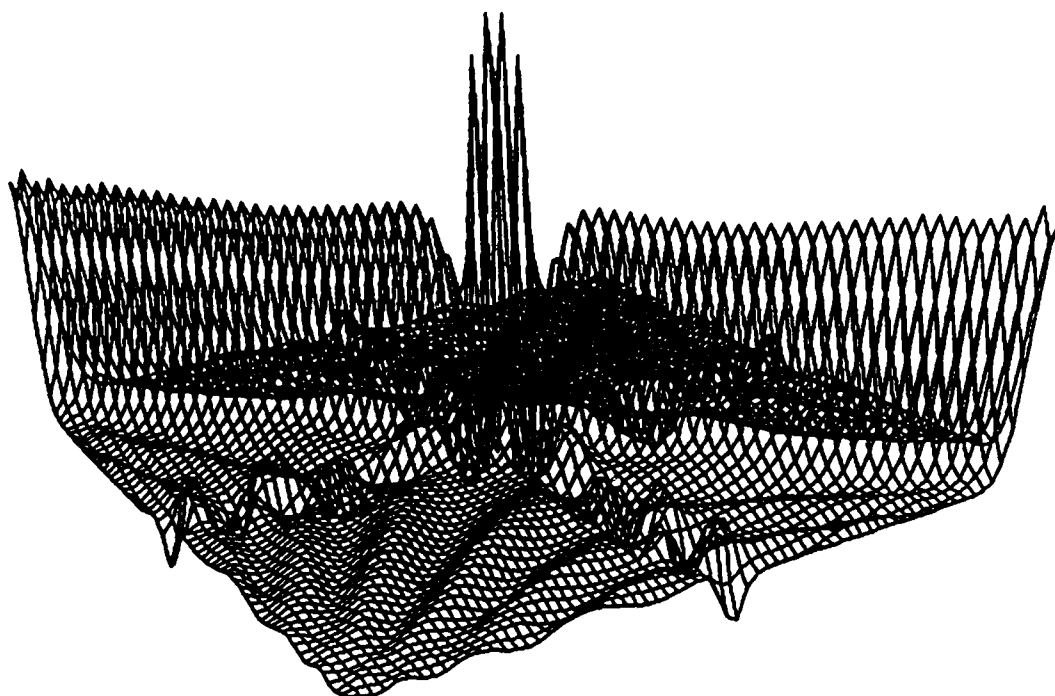
Figure 7h. Re$\rho(x,x',t = 375$ fs$)$.
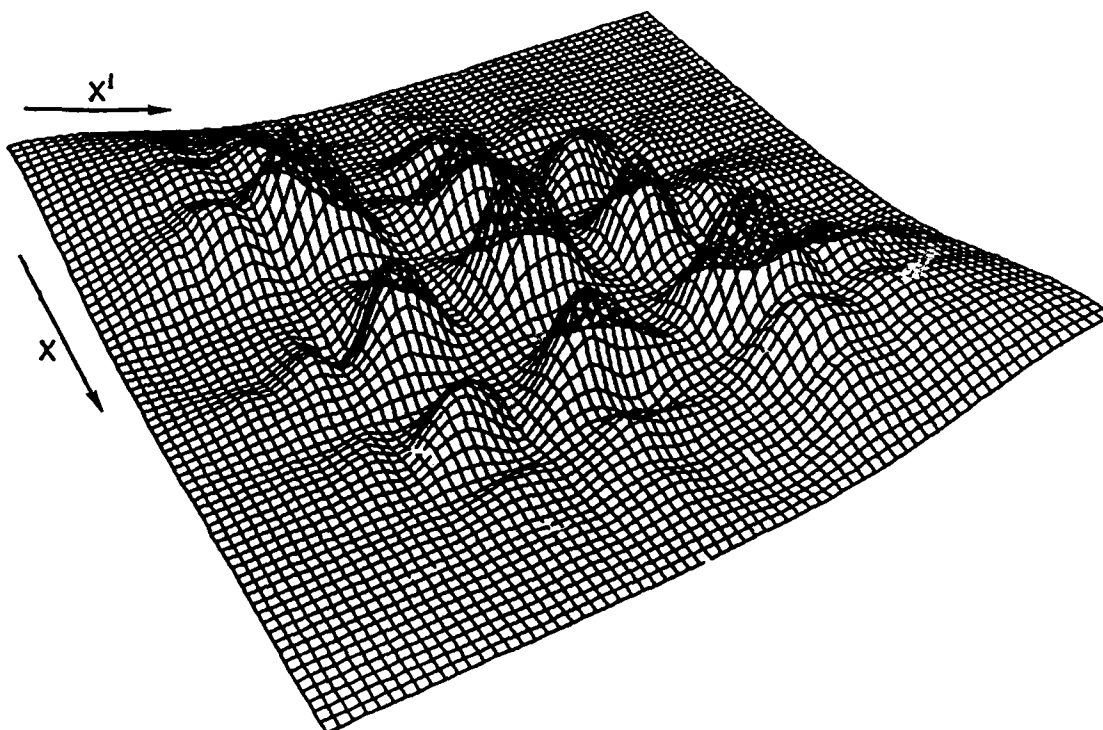
Figure 8a. Reρ(x,x',t) for the double barrier diode at t = 10 fs.
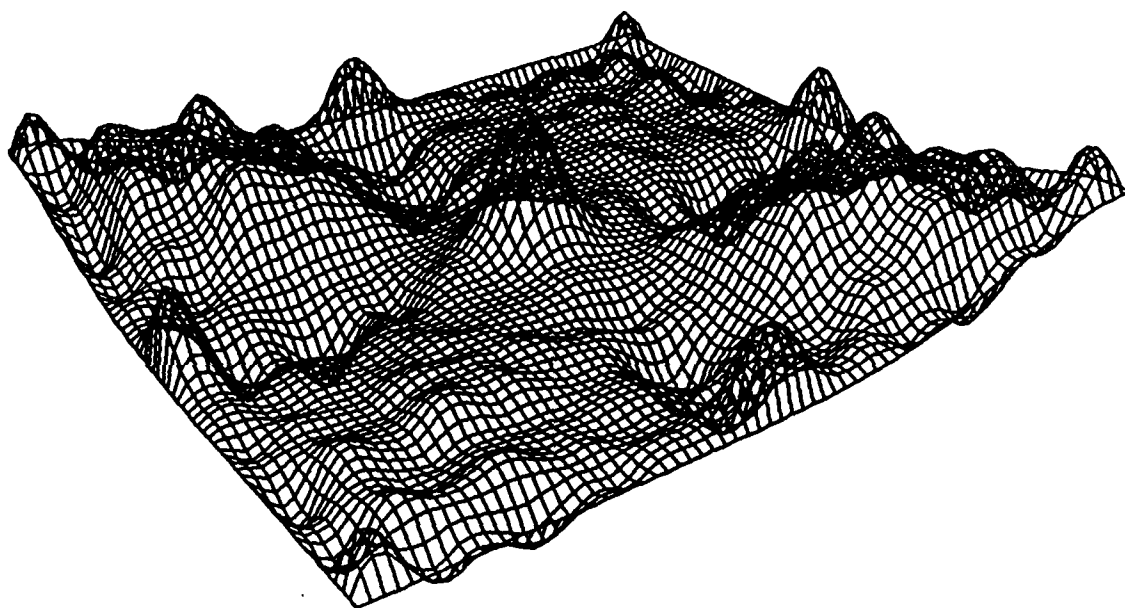
Figure 8b. Rep$(x,x',t = 30$ fs$)$.
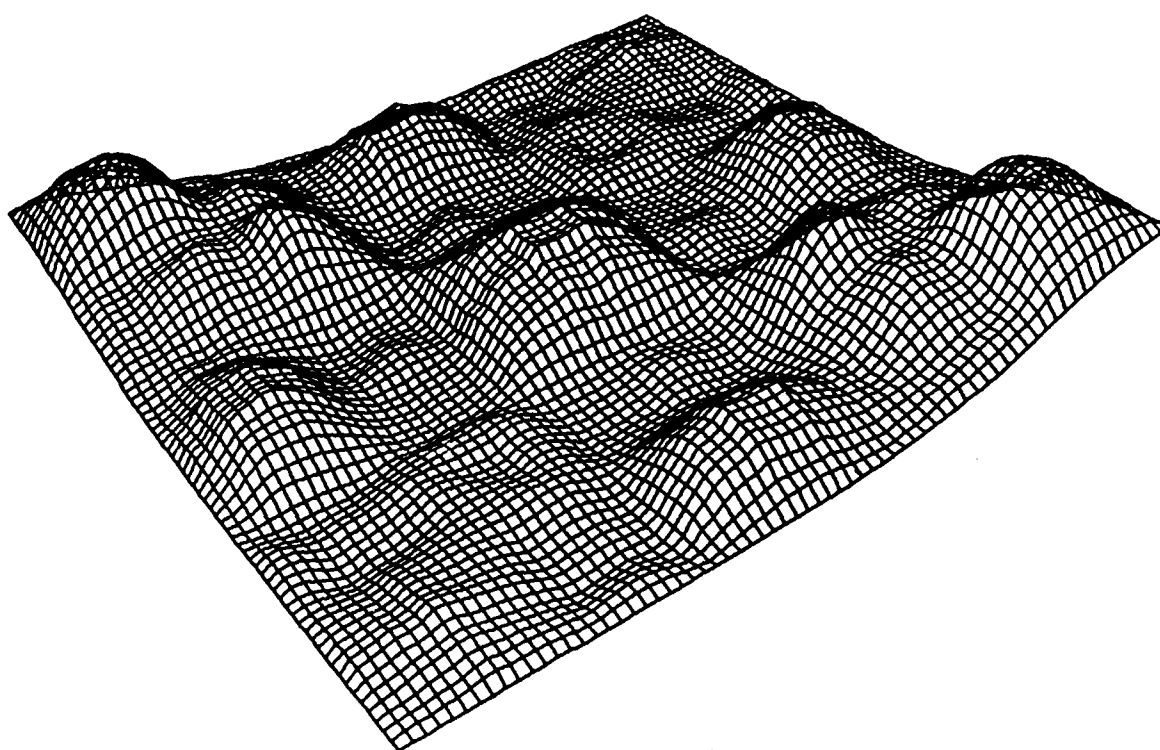
Figure 8c. Reρ(x,x',t = 60 fs).
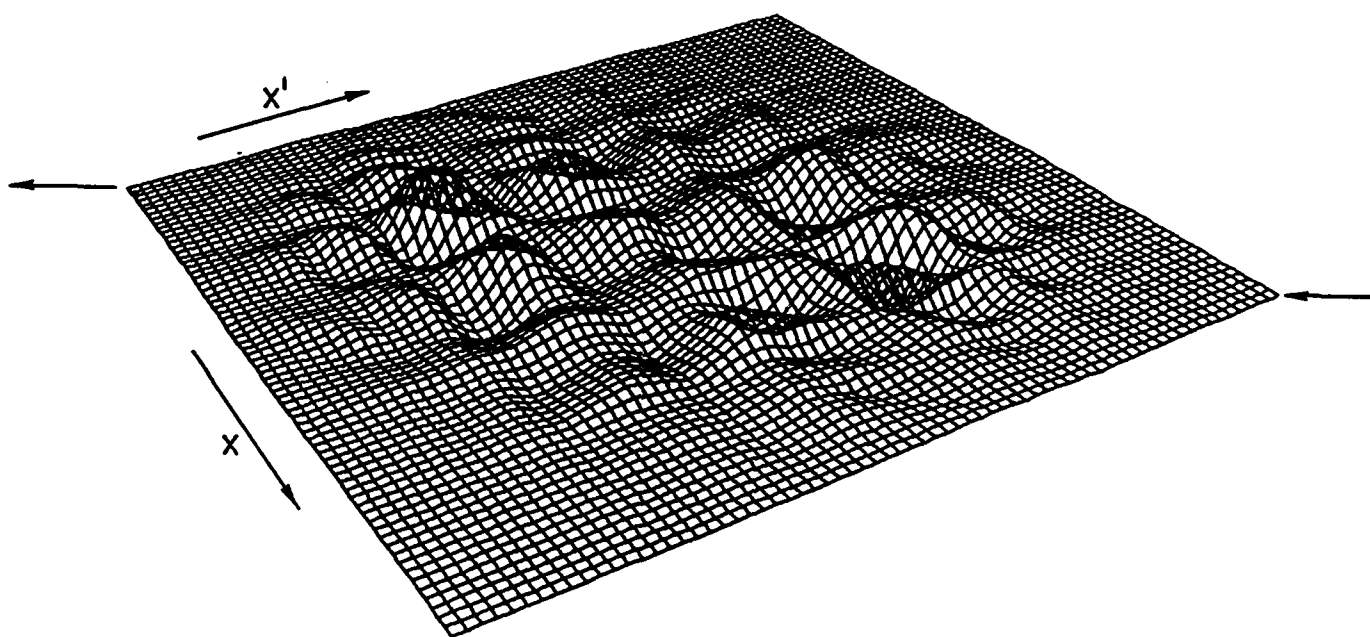
Figure 9a.   Rej(x,x',t) = ρ(x,x',t)V(x,x',t) at t = 10 fs. Positive velocity is in the direction of the arrows.
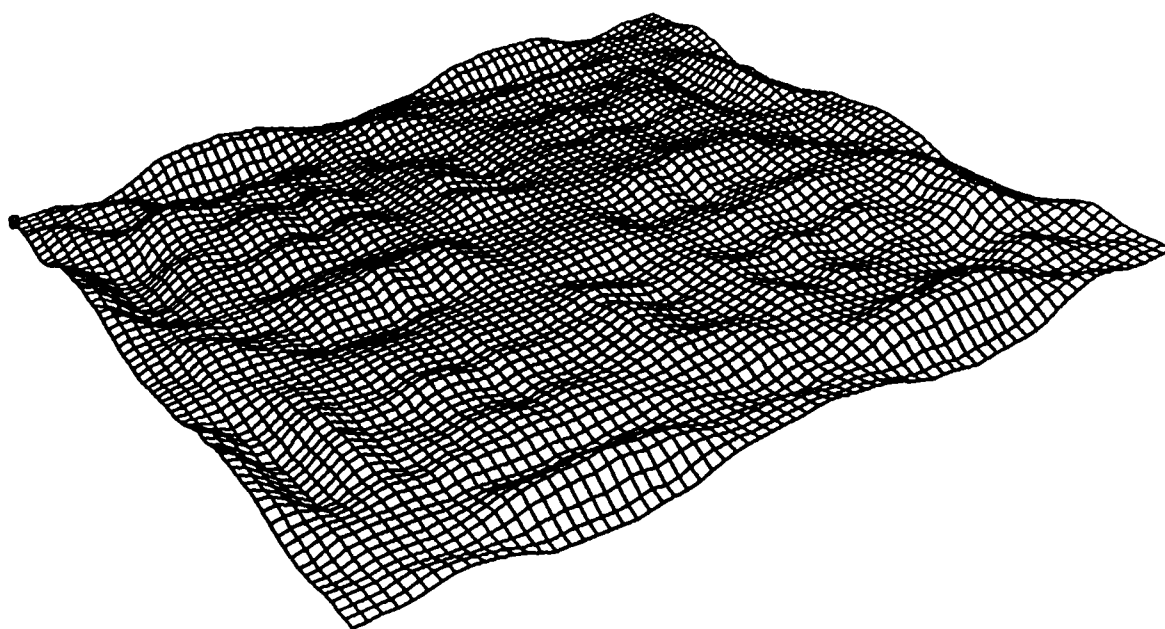
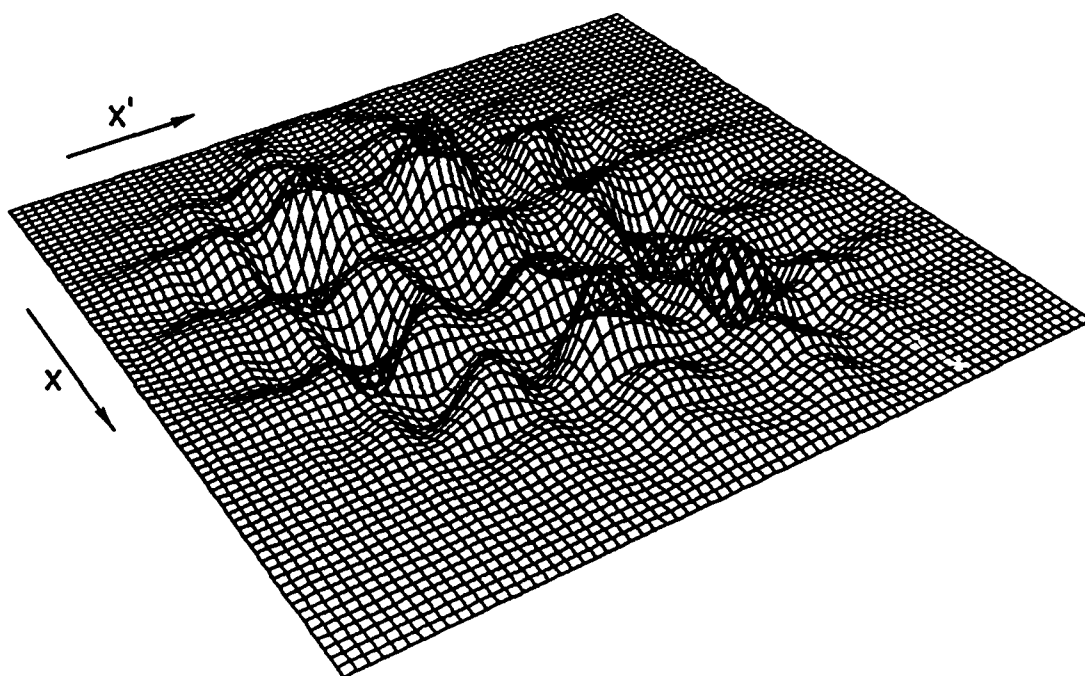Figure 9b.   Rej(x,x',t = 30 fs).
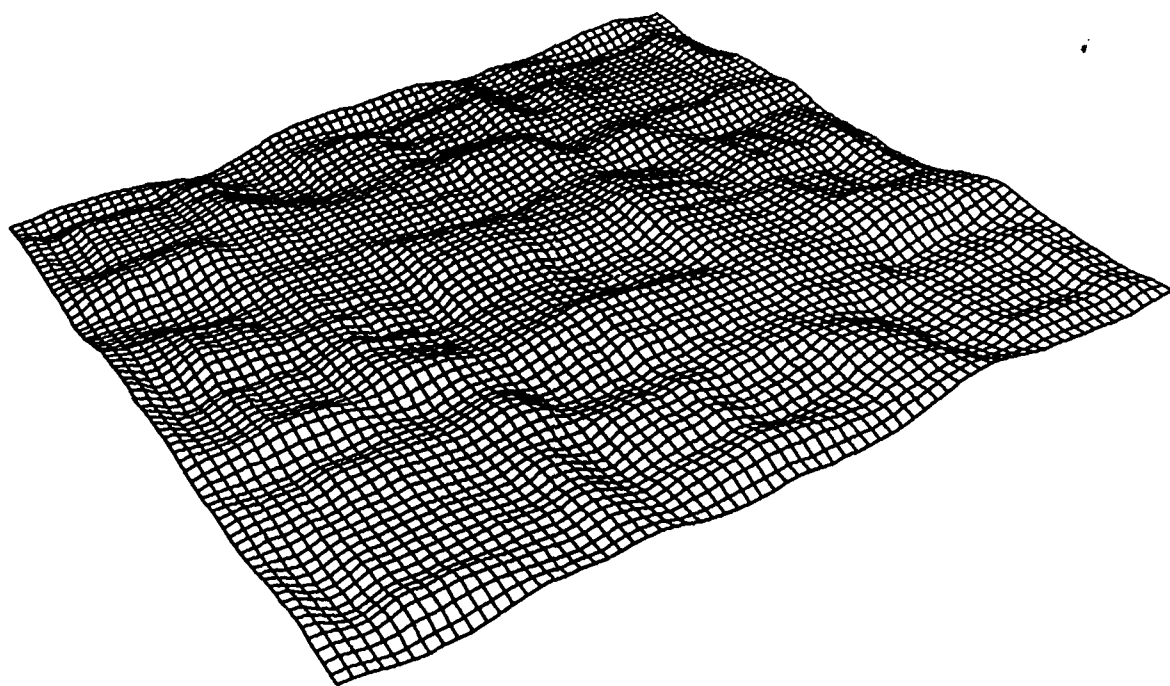
Figure 9c.   Rej(x,x',t = 60 fs).

Figure 10a. Imρ(x,x',t) at t = 10 fs.
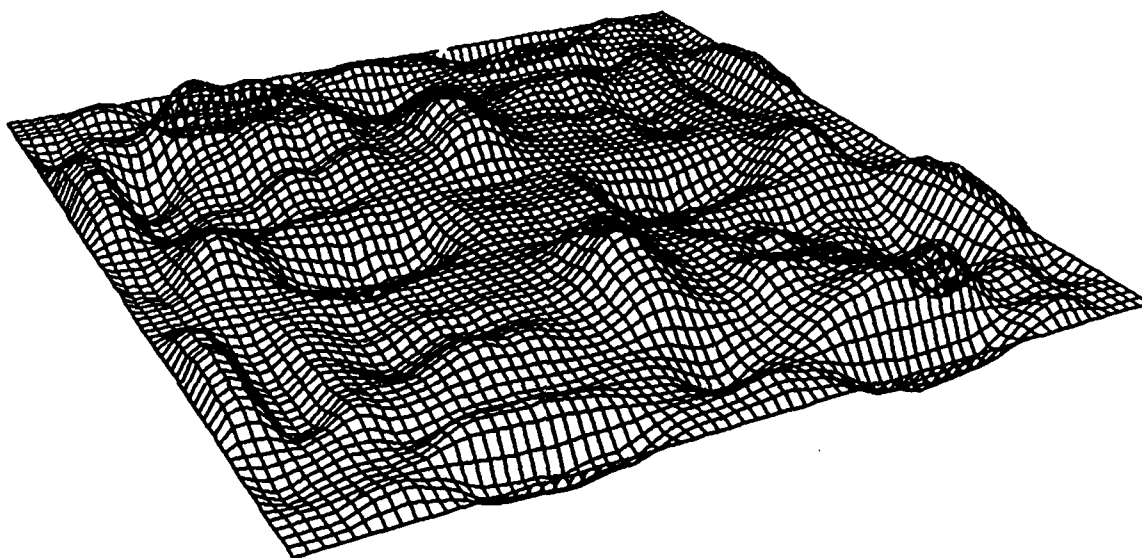
Figure 10b.   Imρ(x,x',t = 30 fs).
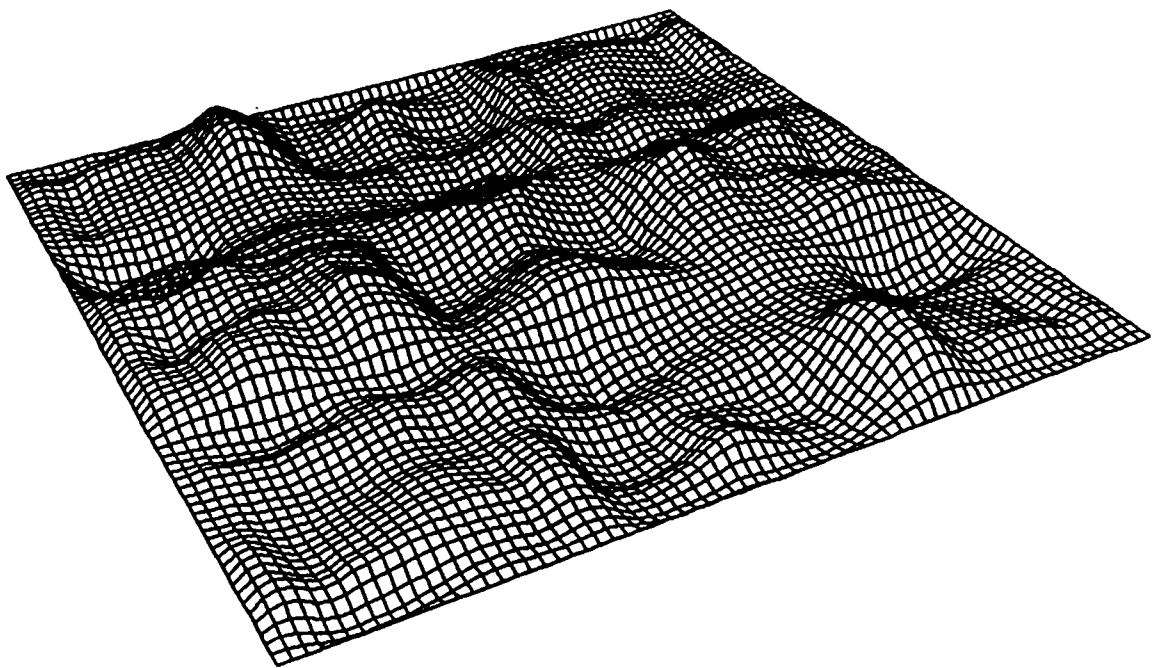
Figure 10c. Imρ(x,x',t = 60 fs).
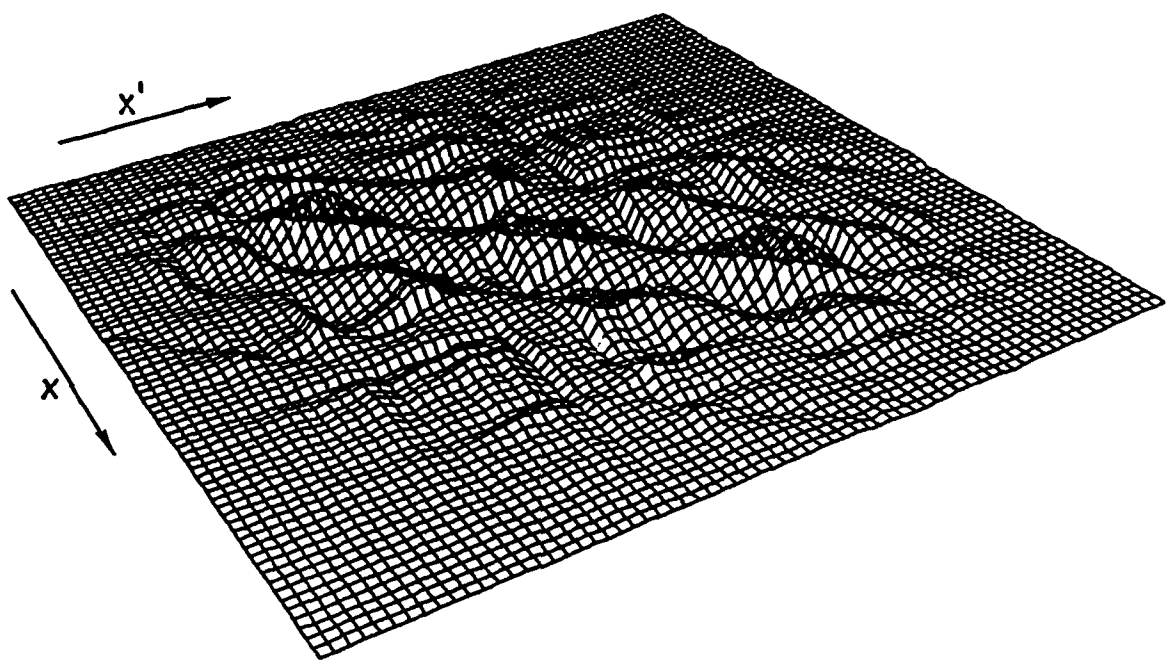
Figure 11a. ReE(x,x',t) at t = 10 fs.
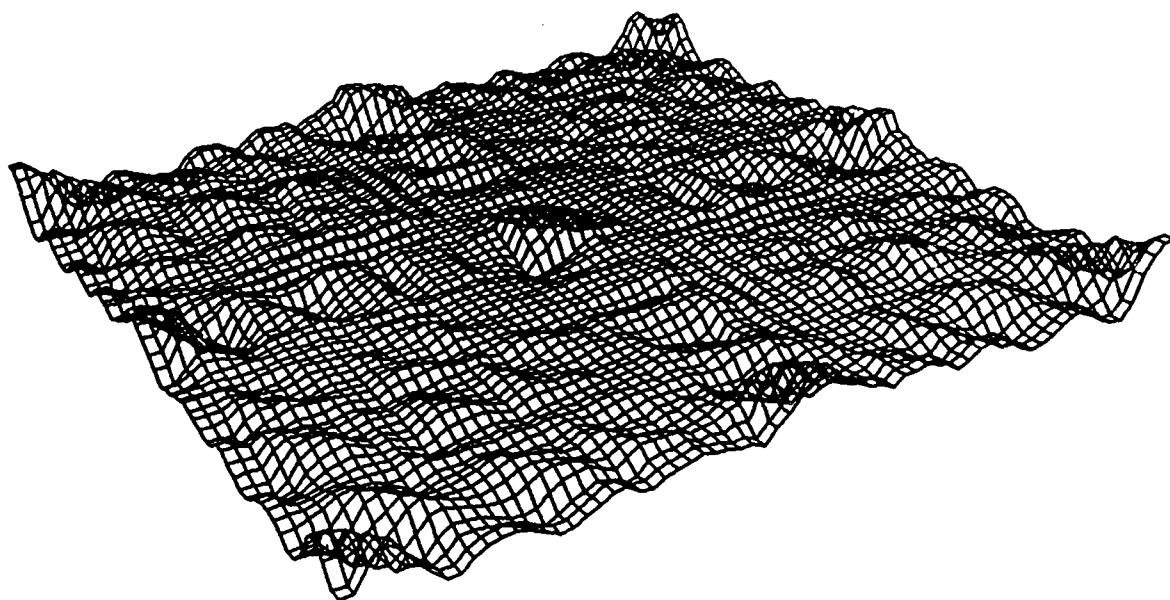
Figure 11b. ReE(x,x',t = 30 fs).
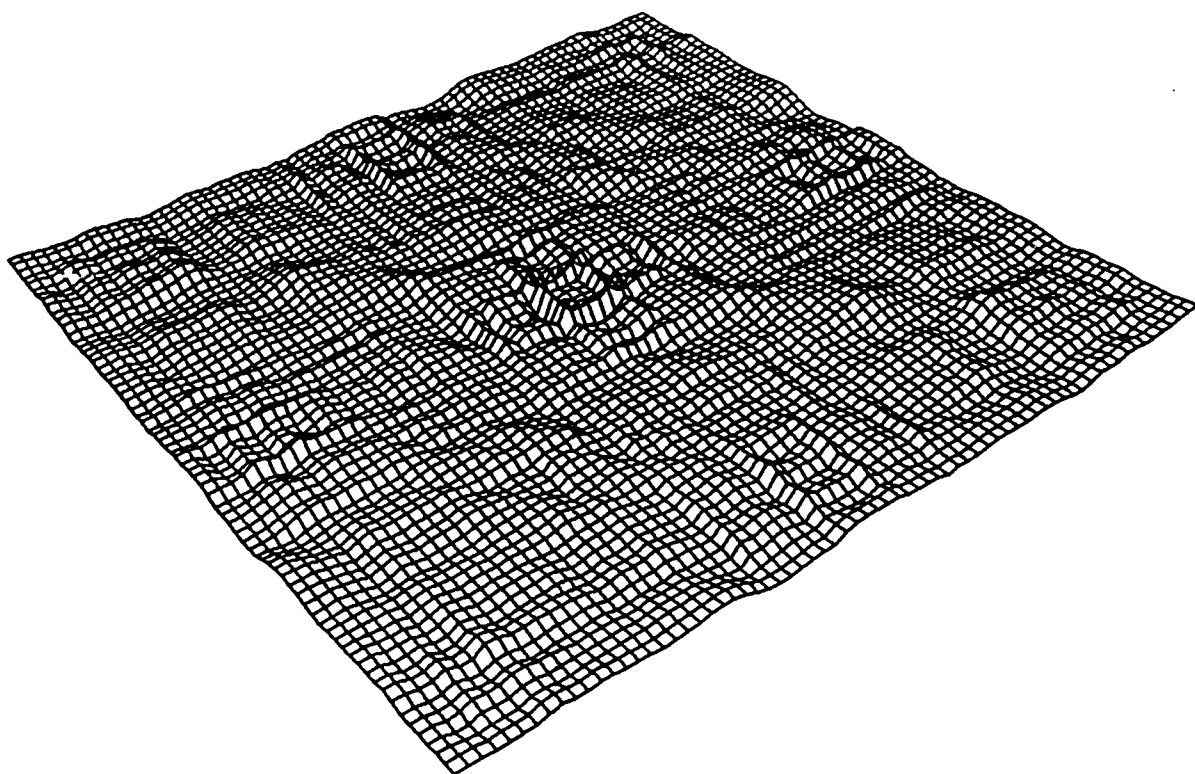
Figure 11c. ReE(x,x',t = 60 fs).

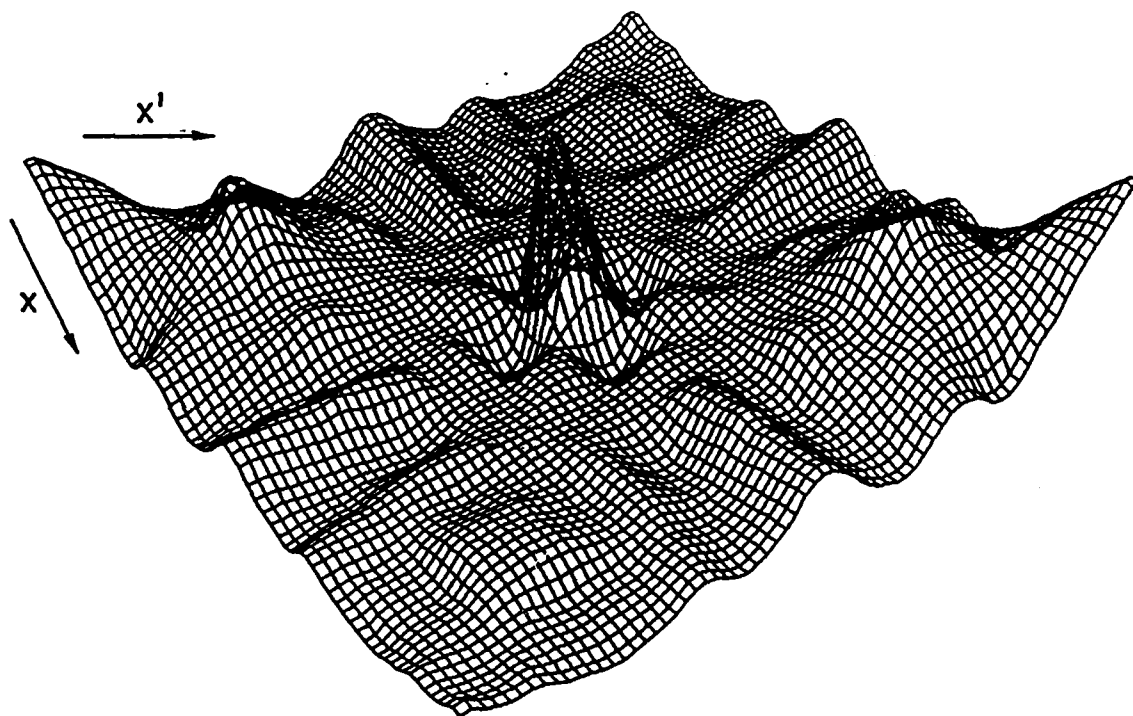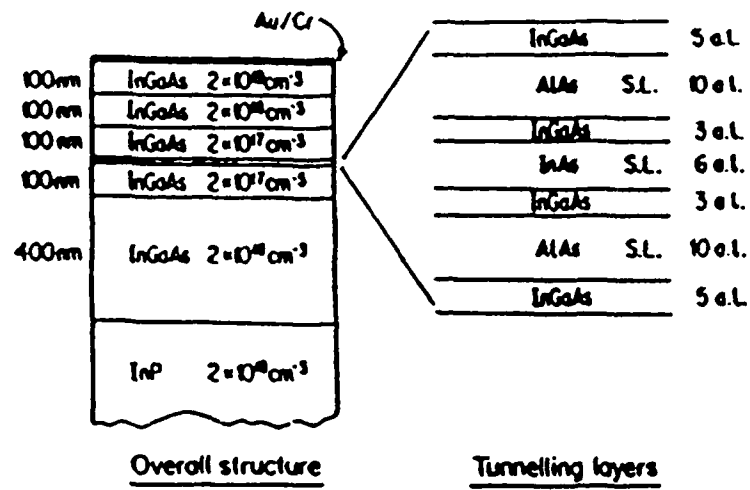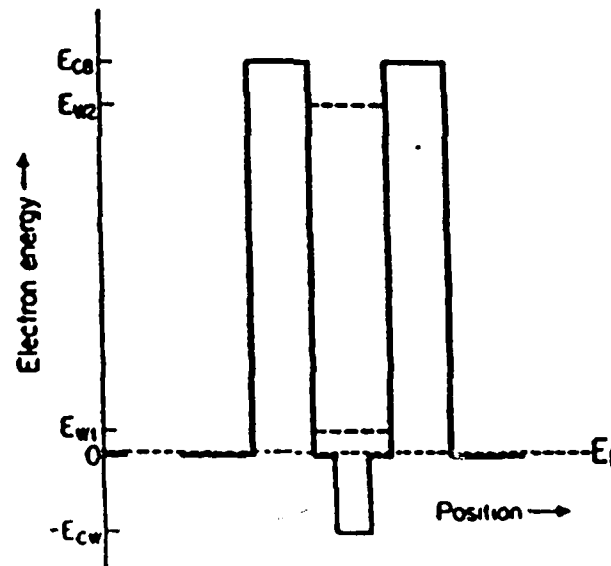Figure 12. Rep(x,x',t) for the structure of Figure 4 with partially transmitting device boundaries.

Figure 13. (a) Schematic cross-sectional layer structure and
(b) the conduction-band diagram of the pseudomorphic
$In_{0.53}Ga_{0.47}As/AlAs/InAs$ double-barrier resonant
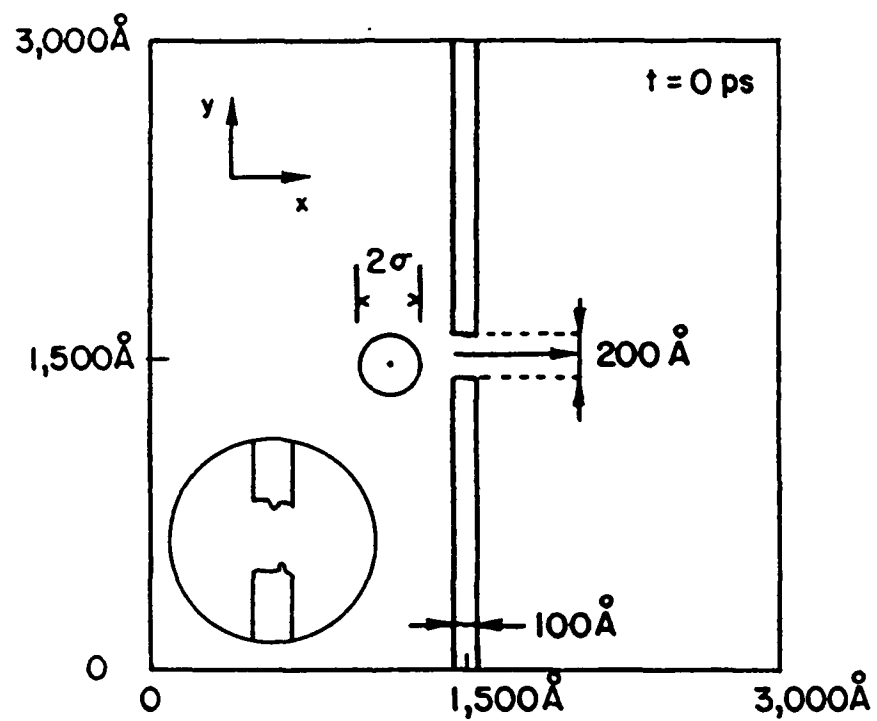tunneling diode grown on InP substrate.

Figure 14.   Schematic representation of electron
diffraction through a narrow slit.

INTRODUCTION TO THE PHYSICS OF GALLIUM ARSENIDE DEVICES

H.L. GRUBIN

SCIENTIFIC RESEARCH ASSOCIATES, INC

P.O. Box 1058

Glastonbury, Connecticut 06033-6058 USA

(203) 659-0333

INTRODUCTION

It is arguable that the history of gallium semiconductor devices, from the early sixties to the present time, fits into three groups. First there was the experimental work of Gunn [1] demonstrating the generation of sustained oscillations upon application of a sufficiently large dc bias. This work opened up the possibility of fabricating bulk microwave and millimeter wave devices, and hastened additional and intense studies of the properties of compound semiconductor devices. Second, there was the study of Ruch [2] whose results suggested that the transient, or non-steady state aspects of semiconductor transport, would improve the speed of devices by almost an order of magnitude. This, of course is the argument behind much of the move toward submicron and ultrasubmicron structures. The third era, is the one we are presently in, and involves the incorporation of gallium arsenide into material engineered highly complex structures, some of which have provided remarkable millimeter wave characteristics, e.g., the pseudo-morphic HEMT [3]. Much of this book is concerned with this third era, and as such we will only briefly touch upon it. Rather, in this chapter we will try to present a road map of the consequences of using compound semiconductors for device applications, using gallium arsenide as the paradigm example.

The band structure of gallium arsenide is familiar to most, and is displayed in Fig. 1. It is a direct band gap material. The minimum in the conduction

band is at Γ with relevant subsidiary conduction band minima at L and X.
The curvature at Γ is such that the effective mass of the Γ-valley
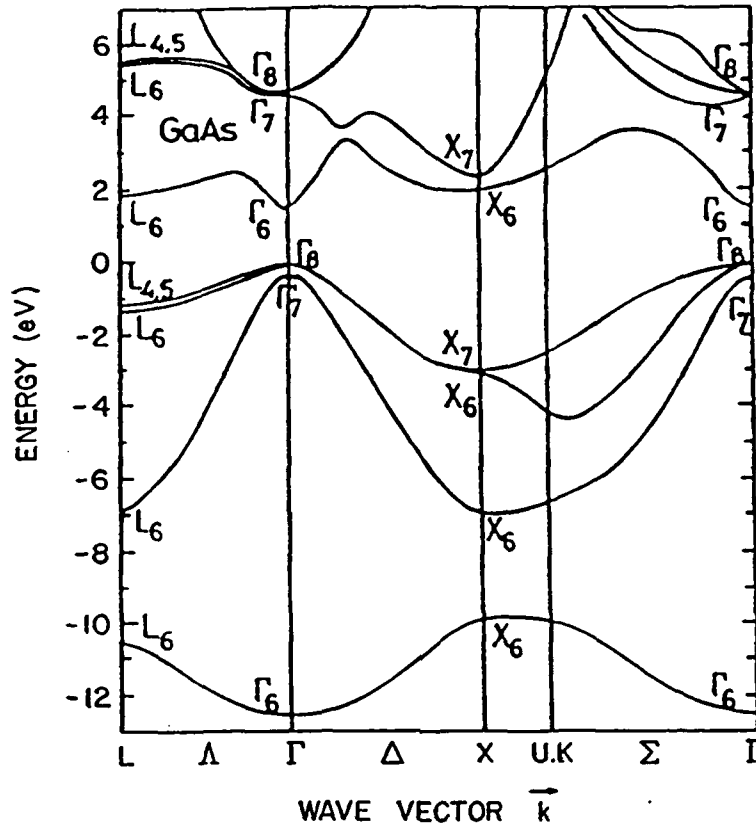
FIGURE 1: Band structure of the semiconductor gallium arsenide [4].

is lower than that of the next two adjacent subsidiary L- and X-valleys. For
the valence band the two valleys of significance are those associated with the
light and heavy holes. We will concentrate on transport contributions from
these five valleys.

In equilibrium, the relative population of electrons in the valleys is
dependent upon the density of available states and the energy separation, e.g.:

$$n_\Gamma^0 = n_L^0 (m_\Gamma/m_L)^{3/2} \exp\Delta/kT \tag{1}$$

where n_Γ and n_L denote the equilibrium density of the Γ- and L-valley

carriers, m_Γ and m_L are their effective masses, and Δ is the Γ–L energy separation. Thus, in equilibrium virtually all of the electrons of interest are in the Γ-valley. For the holes, the valleys are degenerate.

Gallium arsenide is a compound semiconductor. At low values of electric field apart from carrier-carrier scattering, there are three important scattering mechanisms: polar optical phonon scattering, acoustic phonon scattering and impurity scattering. For Γ-valley electrons the contribution to the momentum scattering rate from polar optical phonons is approximately two orders of magnitude larger than that of the acoustic phonon. Since, with regard to mobility, scattering rates are additive, the polar optical phonon is the dominant scatterer. Ideal room temperature electron mobilities are in the range of 8000 to 9000 cm²/vsec. For the subsidiary valleys the effective masses of the carriers are much larger than that of the Γ-valley and the relative contribution of the acoustic phonon increases. Nevertheless, the polar phonon dominates the transport. For holes, the situation is mixed with the dominant scattering being polar and nonpolar deformation potential coupling. For momentum scattering the nonpolar deformation potential scattering dominates.

At high values of electric field and for electrons, nonpolar phonons enter the picture, intervalley transfer from Γ to L takes place, and the situation becomes complex. For example, the spatially uniform, field dependent velocity characteristics of gallium arsenide, ignoring electron-hole interaction displays a region of negative differential mobility, as shown in Fig. 2, where at values of field in excess of 3 kv/cm the mean carrier velocity begins to decrease with increasing electric field. This is an unusual situation and it is perhaps important to recognize that the mean electron velocity of a given species of carrier, assuming a parabolic band, is not decreasing with increasing electric field. Rather the numbers of high mobility electrons are decreasing, due to transfer to the subsidiary larger effective mass valleys.

The situation with holes is different. Here the dominant transport is through the heavy hole. Interband hole scattering is always present even at very low fields, however the relative population is fixed through the ratio of the
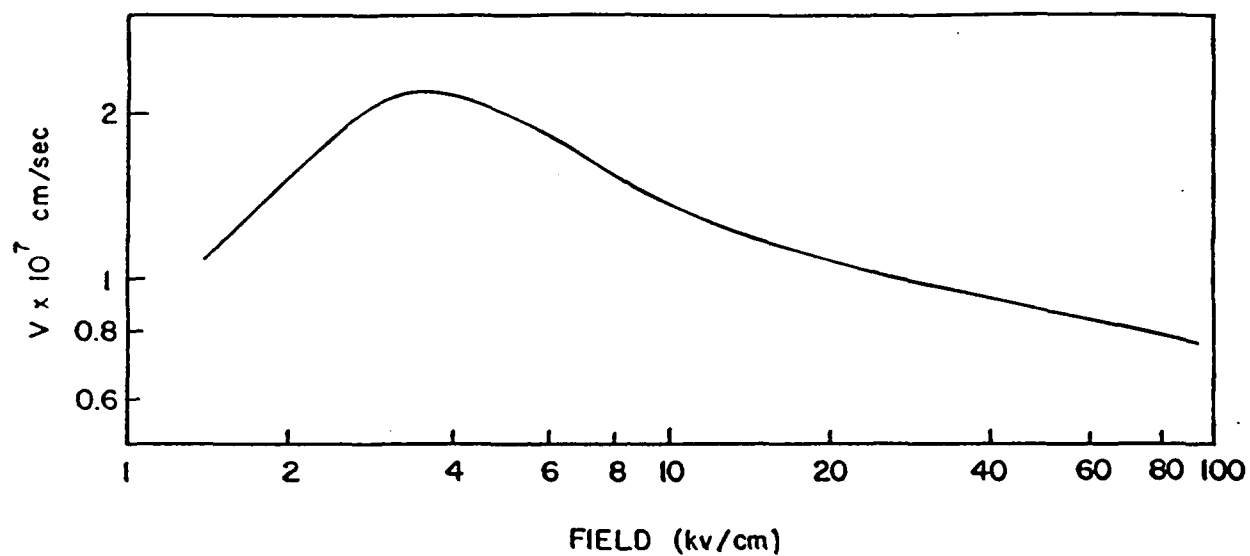
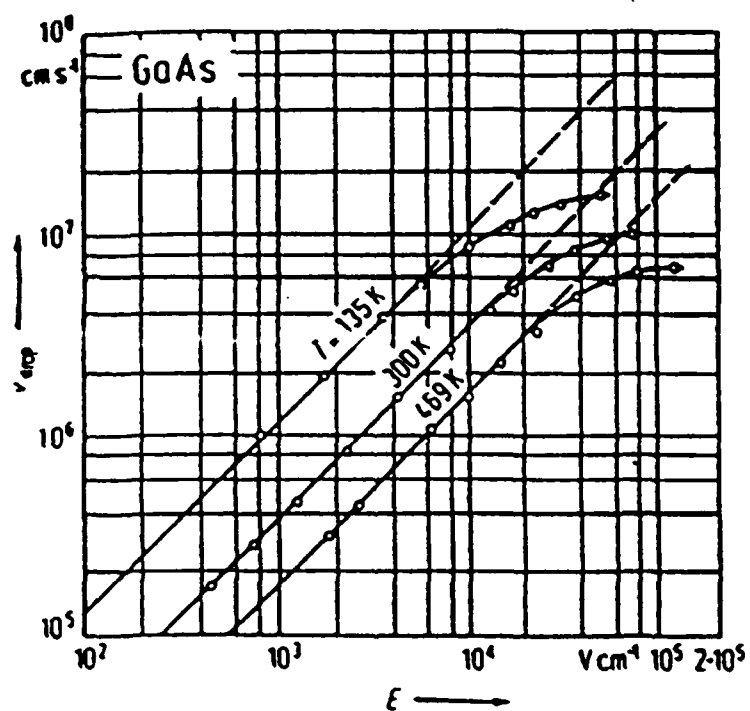FIGURE 2: Field dependent electron mean velocity for gallium arsenide [5].

FIGURE 3: Field dependent hole mean velocity for gallium arsenide [6].

effective masses, and the existence of a dc negative conductance for holes, on the basis of available data is ruled out. The field dependence of the mean hole velocity, ignoring interaction with the electrons, is displayed in Fig. 3, and there are two important features of note: First, there is the extremely low mobility of the holes at low field values. Second, there is the saturated drift velocity, which is expected to be higher than that of electrons at high fields. We note there is no hard data on the high field carrier velocity of holes in gallium arsenide.

Calculations of the type displayed in Figs. 2 and 3 have been described by many workers and are routinely incorporated into simulation codes. Of more recent interest, because of mixed conduction heterostructure devices, are the modifications that may be expected when electron-hole scattering occurs.

O WITHOUT e-h   T = 300°K
● WITH e-h
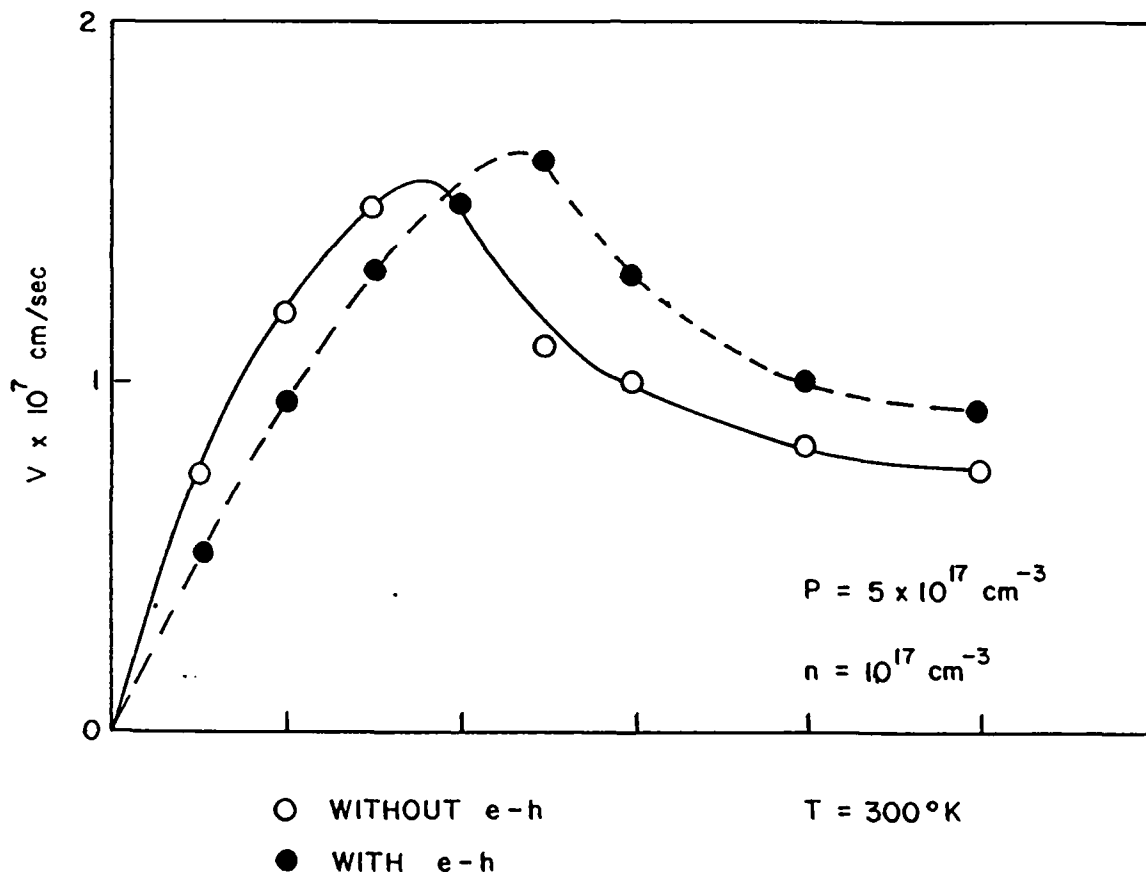
FIGURE 4: Field dependent electron mean velocity for gallium indium arsenide assuming an interaction with heavy holes [7].

However, because of the limited number of studies with GaAs, and because of similarities with other compound semiconductors, results of InGaAs studies are presented. Additionally, because of experimental work on InGaAs [8], the role of carrier-carrier scattering has been most extensively studied for that material. Monte Carlo calculations incorporating electron-hole scattering are displayed in Fig. 4. The results require some detailed discussion and are considered later. Here we simply note that the presence of holes leads to reduction in the low field mobility, but an increased peak carrier velocity. These intriguing results are also anticipated for GaAs.

The question of interest, is how may we expect the role of the complicated compound semiconductor band structure to affect the performance of devices. This is considered next.

THE ROLE OF BAND STRUCTURE ON THE OPERATION OF ELECTRON DEVICES

In examining the role of band structure on the operation of electron devices there are several items of immediate interest: the effective mass, the low field mobility and the direct band gap energy of the binary III-V materials, see Table 1. Additionally, the energy separation of the conduction band

TABLE 1

Critical Parameters of Select Compound Semiconductors [9]

Compound	Effective Mass (multiples of free electron mass at the conduction band minima)	Electron Low Field Mobility (cm^2/V-s)	Direct Energy Bandgap (ev)
GaAs	0.063 (DOS)	9200	1.424
GaP*	$0.25^t/0.91^l$	160	2.78
GaSb	0.042	3750	0.75
InAs	0.0219	33000	0.354
InP	0.079	5370	1.344
InSb	0.0136	77000	0.230
AlAs**	0.71 (DOS)	300	2.98

* The minima in the conduction band are at Δ-axis near zone boundary.
** The minimum in the conduction band is at X.

minima to the subsidiary valleys are listed in Table 2.

TABLE 2 [10] Intervalley Energy Separation			
Compound	ΓL (ev)	ΓX (ev)	LX (ev)
GaAs	0.34	0.48	0.14
GaP	-0.27	-0.39	-0.37
GaSb	0.08	0.37	0.23
InAs	1.27	1.60	0.33
InP	0.63	0.73	0.10
InSb	0.41	0.97	0.56
AlAs	-0.15	-0.79	-0.64

Note that of the seven binary materials listed, five are direct band gap materials, and two, GaP and AlAs are indirect materials. The indirect band gap materials have the highest effective masses of the group and also the lowest mobility. For these materials, GaAs, InAs, InP, InSb, possess regions of negative differential mobility. GaSb, GaP, and AlAs do not. It is perhaps not surprising that the first four mentioned materials possess a region of negative differential mobility, nor that the last two materials do not. In the latter case the minima in energy is associated with a large effective mass, high density of states energy level. The situation with GaSb is peculiar. But here, while the effective mass of the Γ-valley is the smallest of the three, its closeness in energy to that of the subsidiary L-valley is such that at low values of field conduction contributions arise from both the gamma and L-valley, effectively suppressing the contributions of intervalley transfer to negative differential conductivity.

The presence of a region of bulk negative differential mobility has, as a

major consequence, the possibility of electrical instabilities. These instabilities manifest themselves either as large signal dipole dominated oscillations, often referred to as the Gunn effect, or as circuit controlled oscillations, where the semiconductor behaves electrically as a van der Pol oscillator. The binary semiconductors GaAs, InP and InAs have exhibited electrical instabilites associated with bulk negative differential mobility. while InSb has also sustained electrical instabilities, the interpretation of the instability is complicated by the small direct band gap and the possibility of avalanching at low bias levels.

An additional feature of importance is the intrinsic carrier concentrations of some of these materials, as shown in Table 3. It is clear that the intrinsic

TABLE 3 [9]
Intrinsic Concentration

Compound	$n(/cm^3)$
GaAs	2.1×10^6
InAs	1.3×10^{15}
InSb	2.0×10^{16}
InP	1.2×10^8

concentration of InAs and InSb make them unsuitable for a unipolar source. Indeed all transport calculations using these latter materials must necessarily include multi-species transport.

In choosing materials for electron devices, particularly as power sources, a figure of merit has been the peak to saturated drift velocity ratio. From this point of view indium phosphide is an attractive candidate, but this must

be traded with the fact that the low field mobility of InP is less than that of gallium arsenide. A recent study comparing these features suggests that the Γ-valley mobility is the dominant material parameter of submicron structures, whereas the high field saturated drift velocity is the dominant material parameter of micron-length structures [11]. Additionally, if a choice for two terminal sources is to be made between, e.g., InP and GaAs, other issues emerge. For example, the scattering rates in InP indicate a shorter energy relaxation time, than that of GaAs. The consequences of this are higher frequency operation for InP. Thus, at least with respect to these materials the peak to valley ratio of the materials is only one factor in the design of an electrical source.

There is less to say about the effects of negative differential mobility on the operation of avalanche diodes. Here the effects of negative differential mobility conductivity are present but they are overshadowed by the effects of avalanching. For example, recent studies show the presence of domains in the IMPATT simulation whose presence is a direct consequence of negative differential conductivity. These domains can complicate the actual transit time of dipole layers associated with the avalanche generation, but the negative differential mobility is a marginal issue. Such is not the case with three-terminal devices.

For three-terminal device observations of bias dependent white light in GaAs FETs, as from either the drain side of the gate contact and the gate side oi the drain contact are consistent with numerical calculations showing the presence of local high field dipole layers near the gate and drain contacts. In addition, for a range of bias, some devices display a current dropback consistent with bias dependent formation of high field domains and concurrent current oscillations. This last result is shown in Fig. 5 [12]. Remaining questions of interest focus on the manner in which transport in these devices are examined. We begin with the equilibrium description of transport.
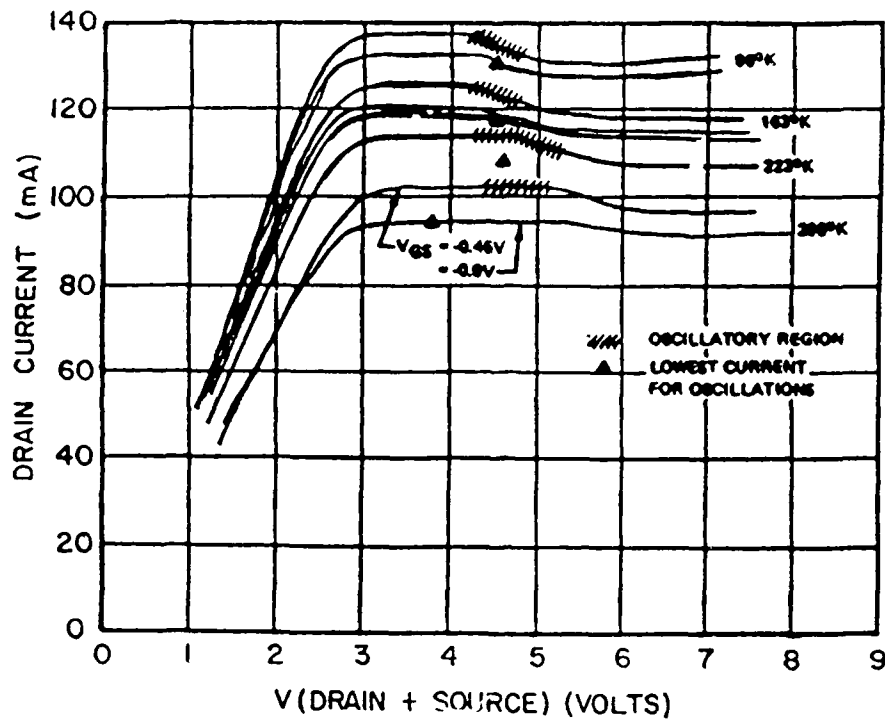
FIGURE 5: Temperature dependent pulsed data for a GaAs FET, with a 3.0μm
gate length, a source to drain separation of 8.5μm, and an
epitaxial thickness of 3000$\pm$ 500 A. Nominal background doping is
10^{17}/cm^3 [12].

EQUILIBRIUM DESCRIPTION OF TRANSPORT

The steady state equilibrium distribution of transport has traditionally
provided most of the description of device behavior. Nevertheless, the
description ignores acceleration. It assumes that the carrier velocity is
determined by the local electric field, and that the total current is governed
by a balance of a drift component and and a diffusive component. Typically
the continuity equation is solved simultaneously with the current equation,
which for electrons is of the form

$$Jn = -e\left[nv_n - D\frac{\partial n}{\partial x}\right] \tag{2}$$

and for holes:

$$J_p - +e\left[pv_p - D_p\frac{\partial p}{\partial x}\right] \qquad (3)$$

Here n and p denote electron and hole concentration, v denotes velocity and D diffusivity. The usual derivations of equations 3 and 4 proceed from a linearization of the Boltzmann transport equation. The assumption is then made that the equation is valid for high field nonlinear transport. Typically the field dependent velocity assumed in these equations is of the type displayed in Fig. (2).

While the use of the field dependent velocity in these equations is universal, the type of diffusivity coefficient used in these studies is almost as numerous as the numbers of workers involved in numerical studies. But a number of important issues are at stake in the description of the diffusivity. For example, if the Einstein relation is used:

$$D - \frac{\mu kT}{e} \qquad (4)$$

then, under equilibrium and/or zero current conditions, the dependence of carrier density on conduction and valence band energy is given by either the equilibrium Boltzmann or Fermi distribution. However, under non-equilibrium conditions (and near zero current conditions), the Einstein relation inadequately describes diffusive transport [13]. To correct for the latter deficiency, the field dependent diffusivity often used in calculations is of a form similar to that shown below [14]

$$D - \frac{\mu kT}{e} + rv_{sat}^2 \qquad (5)$$

where at high values of electric field, the diffusivity only gradually decreases. While the diffusivity coefficient of Eq. (5) more adequately represents high field phenomena, because its field dependence is conceptually consistent only with the assumption of nonequilibrium conditions it is

conceptually inconsistent with equilibrium conditions, and will lead to incorrect built-in potentials [15].

While the drift and diffusion equations clearly offer conceptual difficulties with respect to consistency of physics they nevertheless offer considerable insight into the physics of device operation, and are useful providing we keep in mind it's limitation. For example, instabilities in long GaAs structures, are known to depend critically on conditions at the contacts. A study, in 1969 [16], demonstrated that by experimentally creating different conditions at the boundaries to the active region of GaAs a wide range of different electrical instabilities could be obtained. Corresponding numerical studies were performed through solutions to the above drift and diffusion equations, in which a value for the electric field was specified at the cathode (and anode) boundary. It was found that the boundary dependent electrical behavior could broken into three categories, as summarized in Fig. 6. The key conclusion of the study was that the electrical behavior of compound semiconductors devices was dependent in a detailed way on contact conditions. This same critical result has reappeared numerous times in a variety of different types of structures.

NON-EQUILIBRIUM DESCRIPTIONS OF TRANSPORT

The situation of most interest lies in non-equilibrium transport. Here the most critical area of interest is the incorporation of acceleration into the governing equations.

In examining non-equilibrium transport several approaches have been used. One is the Monte Carlo method, where the trajectory of a particle is followed through its acceleration and subsequent scattering events. In the discussion below results of Monte Carlo calculations will be presented, but we first concentrate on non-equilibrium phenomena as described by the moments of the Boltzmann transport equation. These equations, in their simplest form for parabolic bands, a position dependent conduction band, and a position
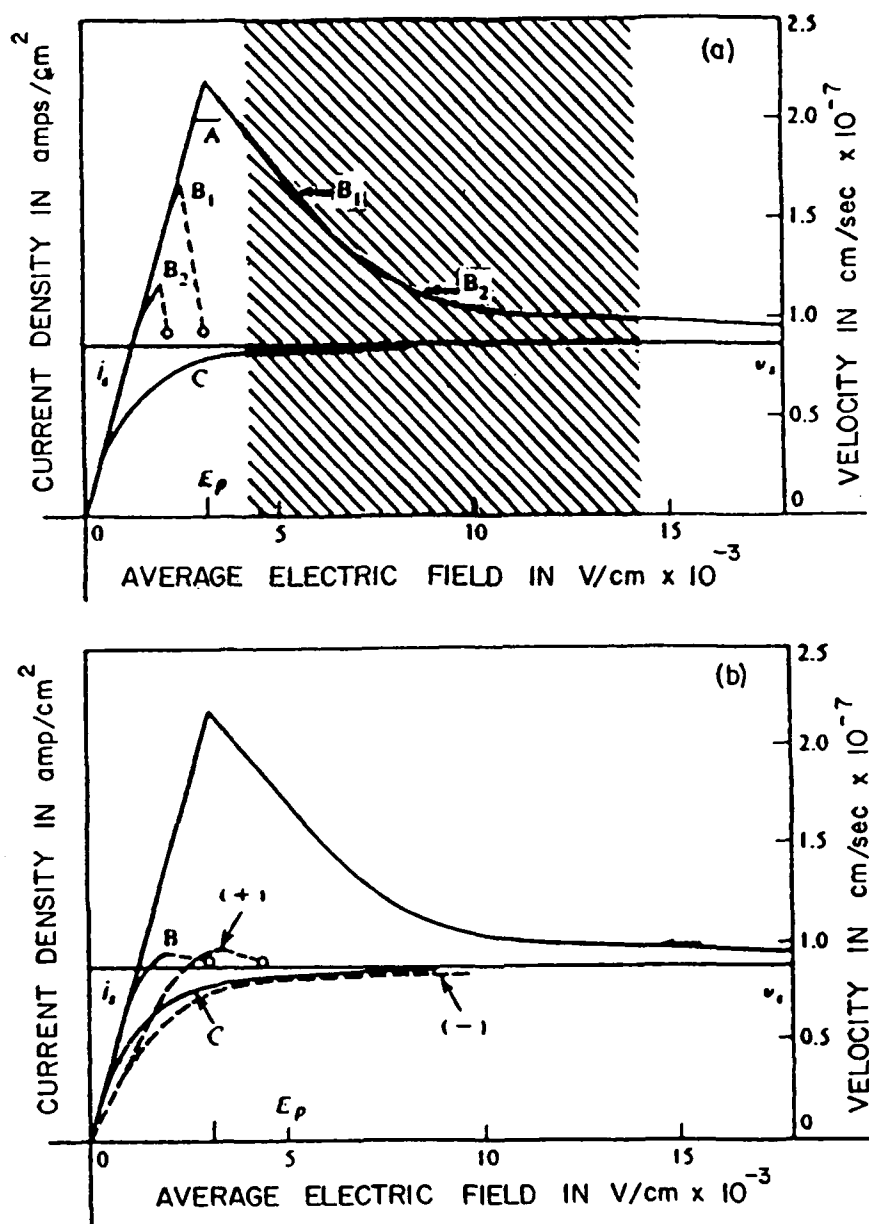
- 13 -

dependent effective mass take the form shown below [17]:

FIGURE 6: a) The v(E) curve and the computer simulated current density j as
a function of average electric field <E> (A through C) for various
cathode boundary fields. The boundary field is zero for curve A,
24 kv/cm for curve C and is indicated by the arrow for curves
B_1 and B_2. The sample is 10^{-2} cm long, has $n-10^{15}$
cm^{-3} and $\mu - 6,860$ cm^2/Vsec. The right- and left-hand
ordinates are related by $j - nev, v_2 - 0.86x10^7$ cm/sec.
b) Experimental $j - <E>$ curves (+) and (−) (dashed) and
theoretical curves B and C (solid). The only significance in the
fact that the low field slopes differ is that the theoretical
curve is for a mobility of 6,860 cm^2/Vsec, whereas the
experimental curve is for a mobility of 4,000 cm^2/Vsec [16].

Carrier balance:

$$\frac{\partial n}{\partial t} + \nabla_r \cdot n\frac{\hbar k_d}{m} - \frac{2}{(2\pi)^3} \int \frac{\partial f}{\partial t}\bigg|_{coll} d^3 k \tag{6}$$

Momentum balance:

$$\frac{\partial}{\partial t} n\hbar k_d + \nabla_r \cdot \frac{(n\hbar^2 k_d k_d)}{m} - -n\nabla_r E_c + qnv \times B - \nabla_r nkT$$
$$+ \left[n\frac{\hbar^2 k_d \cdot k_d}{2m} + \frac{3}{2}nkT\right]\frac{\nabla_r m}{m} + \frac{2}{(2\pi)^3} \int \frac{\partial f}{\partial t}\bigg|_{coll} \hbar k d^3 k \tag{7}$$

Energy balance:

$$\frac{\partial}{\partial t}\left[n\left(\frac{\hbar^2 k_d \cdot k_d}{2m} + \frac{3}{2}kT\right)\right] + \nabla_r \cdot nv\left(\frac{\hbar^2 k_d \cdot k_d}{2m} + \frac{5}{2}kT\right)$$
$$- -nv\cdot\nabla_r E_c + \frac{2}{(2\pi)^3} \int \frac{\partial f}{\partial t}\bigg|_{coll} \frac{\hbar^2 k \cdot k}{2m} d^3 k \tag{8}$$

In the above $\hbar k_d$ is the mean momentum of the carriers, T is the carrier temperature and for electrons, E_c is the position dependent conduction band energy, and B an applied magnetic field. The terms on the right side represent scattering and/or electron-hole interaction, as through avalanching. For example, the right side of Eq. (21) represents intervalley scattering. If avalanching occurs, generation is expressed through an energy dependent ionization coefficient [18]. If a carrier temperature model is assumed, then carrier generation is given by

$$n\alpha(T) \tag{9}$$

where $\alpha(T)$ is the ionization state. In the absence of a first principle determination of $\alpha(T)$, the following relation can be assumed as a starting point

$$\alpha(T) - \alpha^*(F)v(F) \tag{10}$$

where $\alpha^*(F)$ and $v(F)$ are the equilibrium ionization rates, and field
dependent velocities, and the relation between T and F is determined from the
equilibrium solution. While the Eq. (10) relation is uncertain, it has the
conceptual advantage of relating ionization to energy, rather than field.

But, perhaps the most significant feature of these equations is the presence
of acceleration, both spatial and temporal in the momentum balance equation.
These acceleration terms are absent from the drift and diffusion equations.
Additionally, under equilibrium conditions, and hence, zero current (i.e.
$n\hbar k_d/m = 0$), the electron temperature model teaches that for any spatially
nonuniform structure, such as a pn junction, the electron temperature   is
everywhere constant and equal to the ambient. Thus, conceptual problems
arising from the form of the diffusion contribution to the drift and diffusion
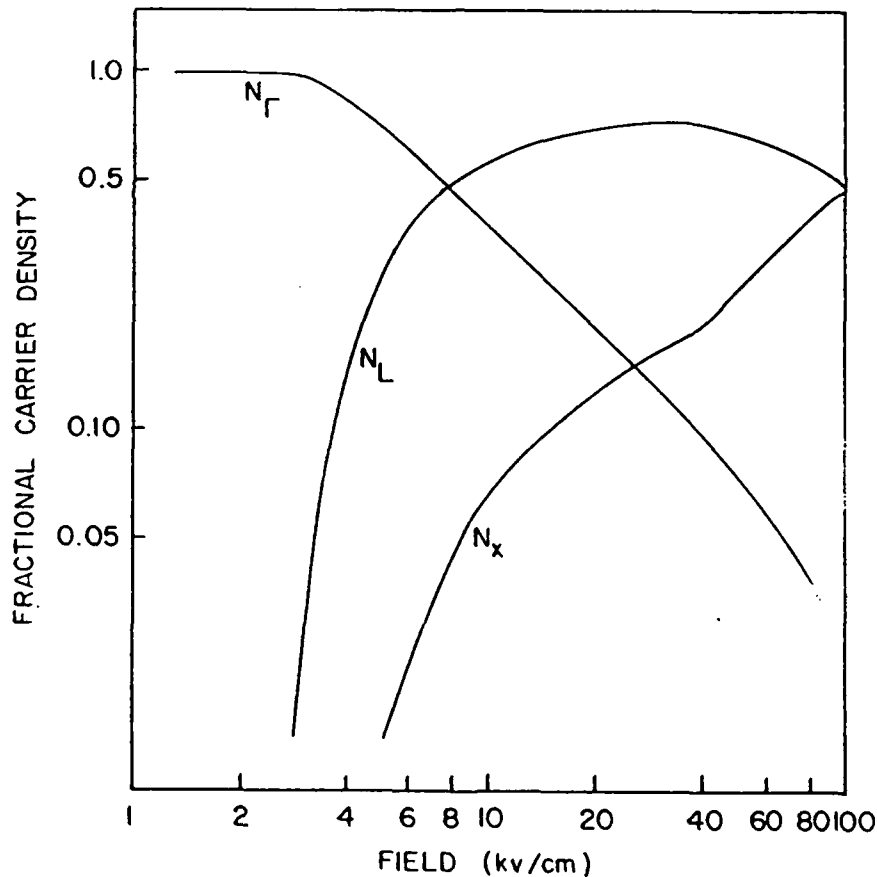equations do not enter here. Note that a generalized drift and diffusion

FIGURE 7:   Steady state uniform field carrier distribution for Γ-L-X
            orientation in GaAs [5].

current term is obtained when the left side of Eq. (7) is set to zero.

Equations of the type shown above provide considerable amount of information with regard to transport. For example, with a Γ-L-X orientation in GaAs the distribution of carriers as a function of field is shown in Fig. 7. Here the relative distribution of carriers in each of the valleys is determined by the distribution of temperature in each of the values, which in turn is driven by the electric field, as shown in Fig. 8. Note that for set fields below 4Kv/cm. The carriers reside in the Γ-valley. At fields above 6Kv/cm, the L-valley population exceeds that of the Γ-valley. It should be emphasized, however, that because of the very low subsidiary valley mobility, most of the current, is for fields up to 50 Kv/cm, is carried by the Γ-valley carriers.
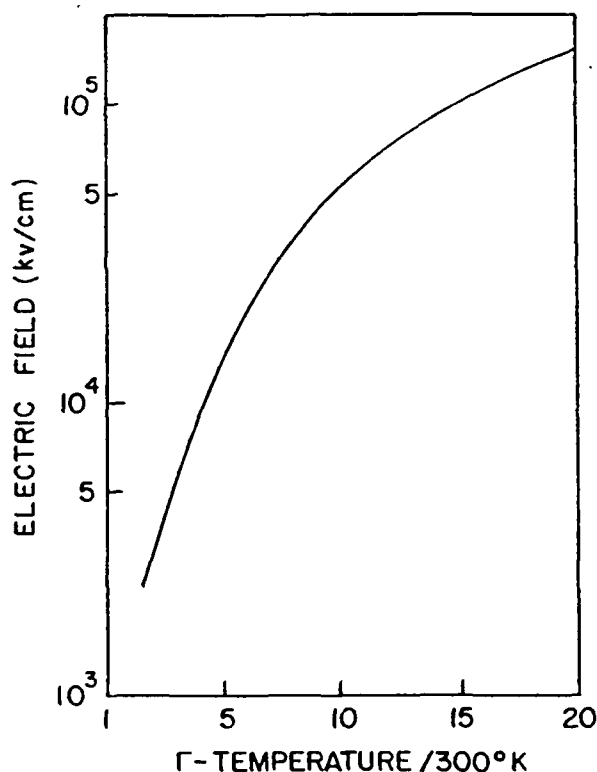

FIGURE 8: Temperature dependence of electric field for a Γ-L-X orientation in GaAs. Electric field is the independent variable [5].

The interest in nonequilibrium equations lies not in the steady state uniform field distribution, but in transients and nonuniform fields. The transient

distribution of carrier density and velocity for electrons subject to a sudden change in electric field is shown in Fig. 9 where we note the high peak velocity.

The high peak velocity in Fig. 9a is primarily associated with Γ-valley

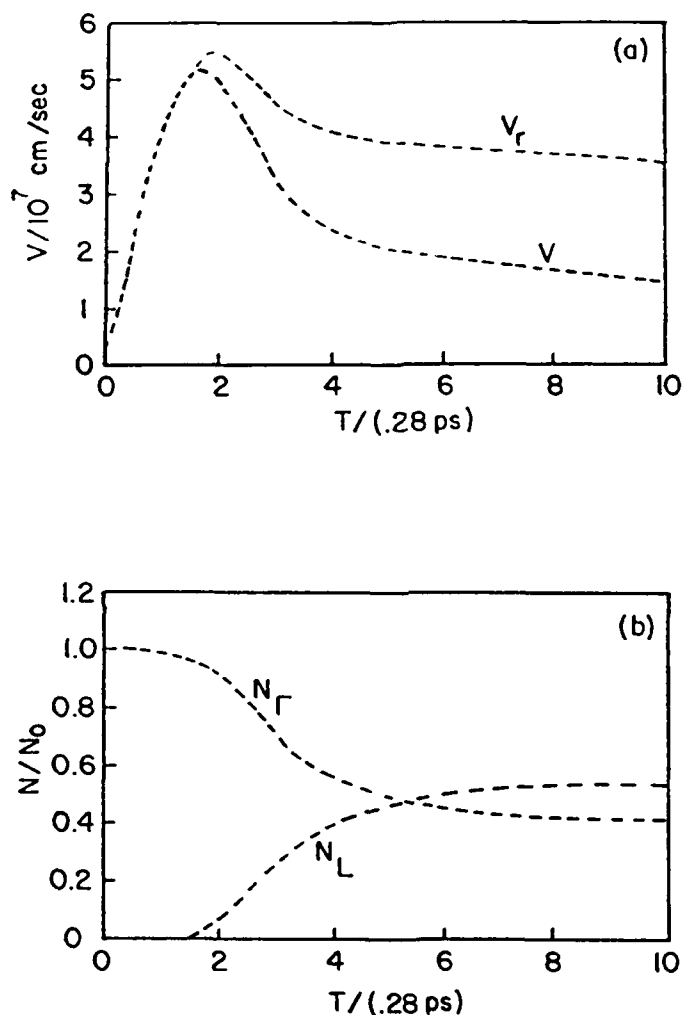FIGURE 9: Transient overshoot for a field of 9.6 Kv/cm [5]. Γ-L-X orientation, and an applied field of (a) Γ and mean velocity, (b) Γ- and L-valley carrier density.

transport. Indeed the Γ-valley velocity at 2.8 ps is near 4×10^7 cm/sec.

The mean velocity

$$v - (n_\Gamma v_\Gamma + n_L v_L + n_X v_X)/N_o \qquad (11)$$

where n denotes net population of the Γ, L and X levels, and N_o is the total carrier density, is also shown in Fig. 9. We note that the significant drop in mean velocity is a consequence of electron transfer from the central to the satellite valleys (see Fig. 9b). Also note a transient decrease in Γ-valley velocity. This is a consequence of the difference between the energy (longer) and momentum (shorter) relaxation times in GaAs. The time independent spatially nonuniform situation also displays overshoot effects.

The situation when, in one dimension space charge effects are introduced is displayed in Figs. (10) and (11), where for a gallium arsenide device of different lengths we show the field and space charge distribution of the
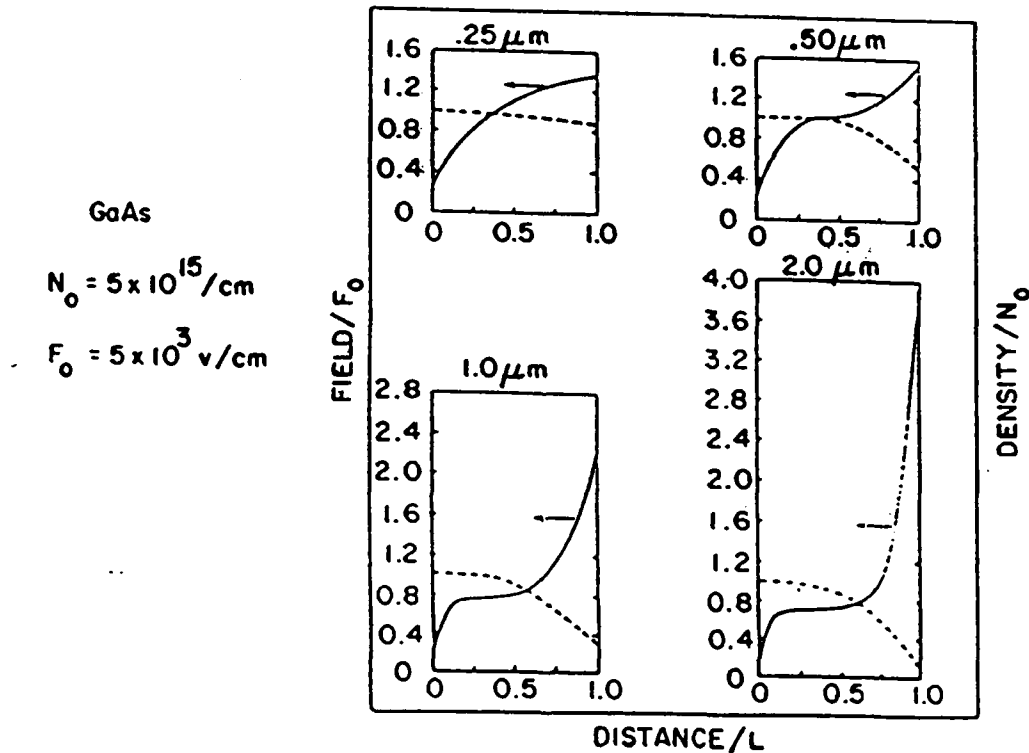

FIGURE 10: Electric field and Γ-valley carrier distribution for GaAs (with a two-level model) two-terminal devices of different lengths and a mean field of 5 Kv/cm [5].
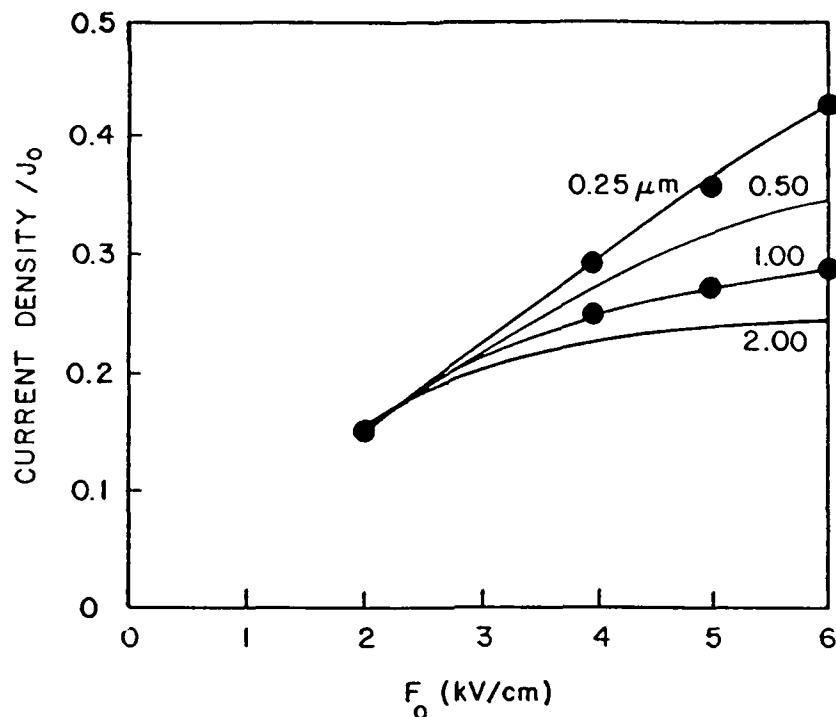
FIGURE 11: Current-voltage characteristics, as a function of length, for the structures of Fig. 10 [5].

Γ-valley electrons and the current voltage characteristics . The feature to note is that as the device length decreases the current drive increases. Note: in all cases the field is nonuniform and increases toward the anode. Electron transfer exists for all four structures with the greatest amount of transfer occurring for the longest device. Additionally, since high field values are synonymous with carrier accumulation, we see that electron transfer here is accompanied by an accumulation of L-valley carriers. Saturation in the current density occurs at high bias, and even for the shortest device there is electron transfer at the anode side of the device.

The role of non-equilibrium transport on two-dimensional simulations is discussed for the vertical three terminal GaAs permeable base transistor [19],

- 20 -

one cell of which is displayed in Fig. 12. One important advantage of this structure is the parallel placement of the source and drain contacts and the absence of any substrate through which current can flow and reduce the

EMITTER
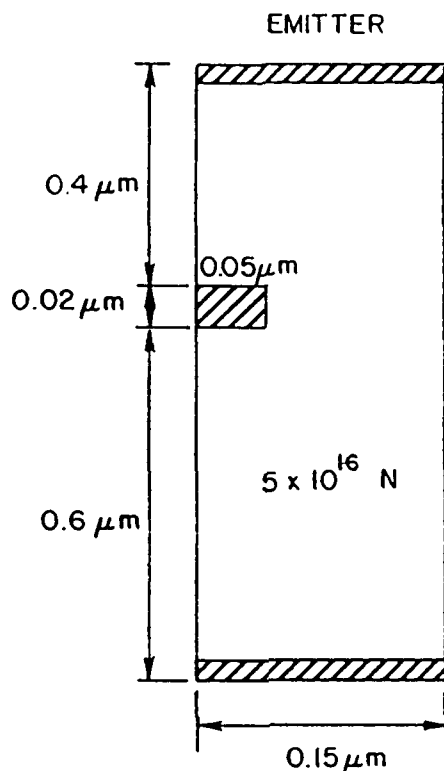
FIGURE 12: Dimensions and doping level of the simulated PBT [20].

transfer characteristics of the device. The simulations were performed for a one micron long source to drain region, and a 200 angstrom gate. Also for comparison we gave included results of the drift and diffusion equation simulations.

The computed I-V characteristics of the device shown in Fig. 12 are presented in Fig. 13. The moment equation results are extrapolated to the origin, as indicated by the long broken lines. The shorter dashed curves show the results for the DDE. The comparison shows that the predicted current levels are significantly higher for the moment equation solutions, a result consistent with FET calculations performed by Cook and Frey [21] who used a

- 21 -
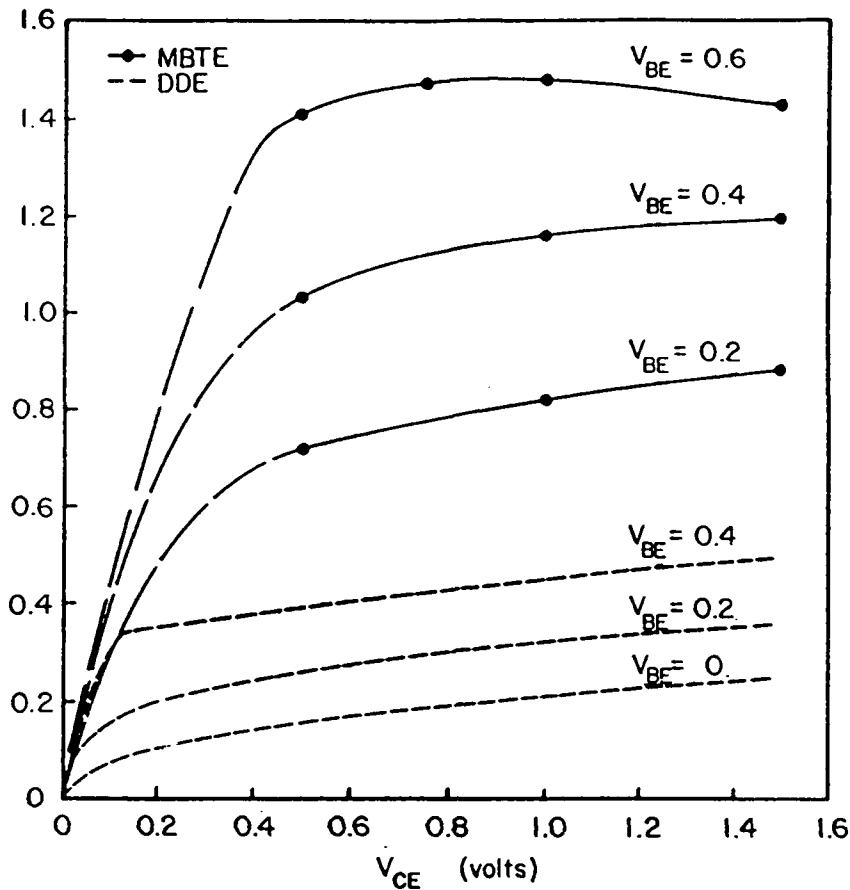
FIGURE 13: Collector current vs. collector emitter voltage for different
values of base emitter voltage [20].

highly simplified momentum-energy transport model. The present calculation
results also indicate a region of negative differential forward conductivity
at V_{BE} - 0.6V. The origin of this phenomena a consequence of electron
transfer. The presence of a dc negative forward conductance is also a feature
of PBT measurements, but is clearly absent from DDE simulations [20].

A comparison between the total carrier density distribution along the center
of the channel for drift and diffusion and MBTE solutions is shown in Fig. 14
for V_{CE} - 1.0V and V_{BE} - 0.4V. The moment equation prediction for the
Γ-valley carrier density is also shown.

As seen in Fig. 14 for the DDE simulations, the carrier density reaches a maximum between base contacts and displays a significant dipole layer. Here,
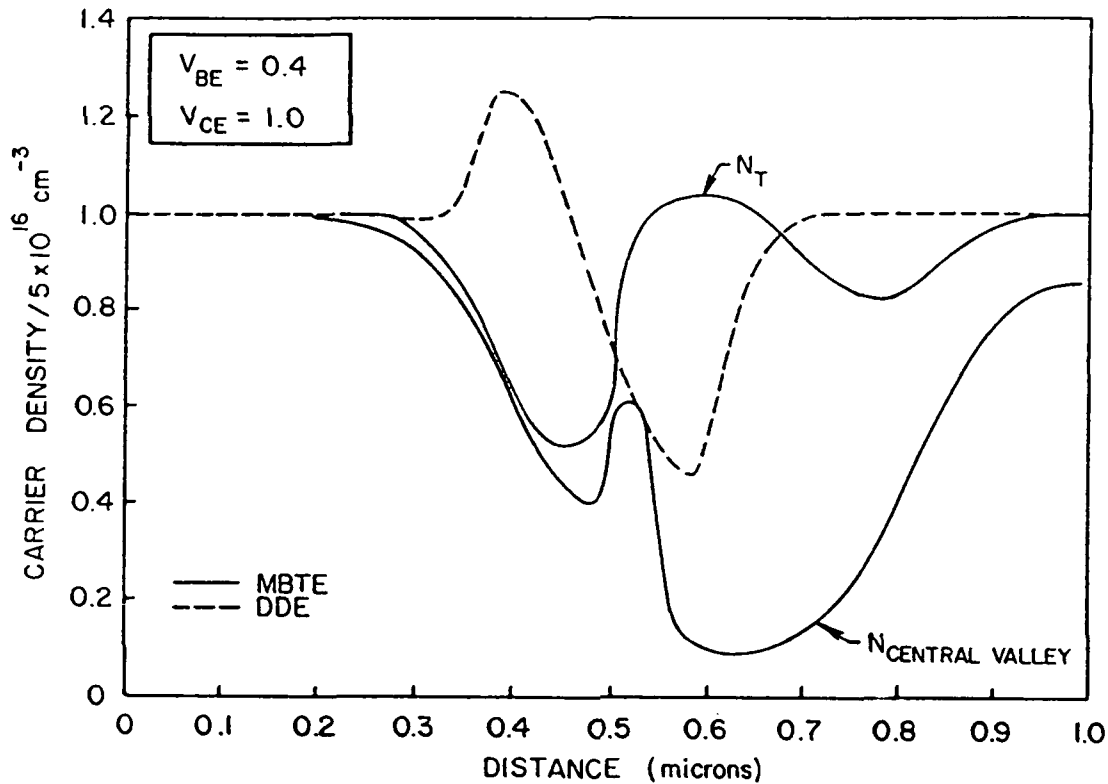
FIGURE 14:   Carrier density vs. distance along center of channel for the PBT [20].

with the velocity in saturation and the cross sectional area at a minimum, the carrier density must increase to maintain current continuity.   In the moment equation simulation the constraints of current continuity are more complex. First a decrease in the cross sectional areas is, as in the DDE, accompanied by an increase in field along the channel.   The field increase under both equilibrium and non-equilibrium conditions is qualitatively similar, as may be

observed from Fig. 15 which shows the potential distribution along the center
of the PBT channel. However, consequent changes in electron temperature, both

FIGURE 15: Potential vs. distance along center of channel for the PBT [20].

increasing and decreasing, lag behind the equilibrium state. This leads to
velocity overshoot (see Fig. 16) and a delay in electron transfer. As a
result, for nearly the first half of the device transport is almost
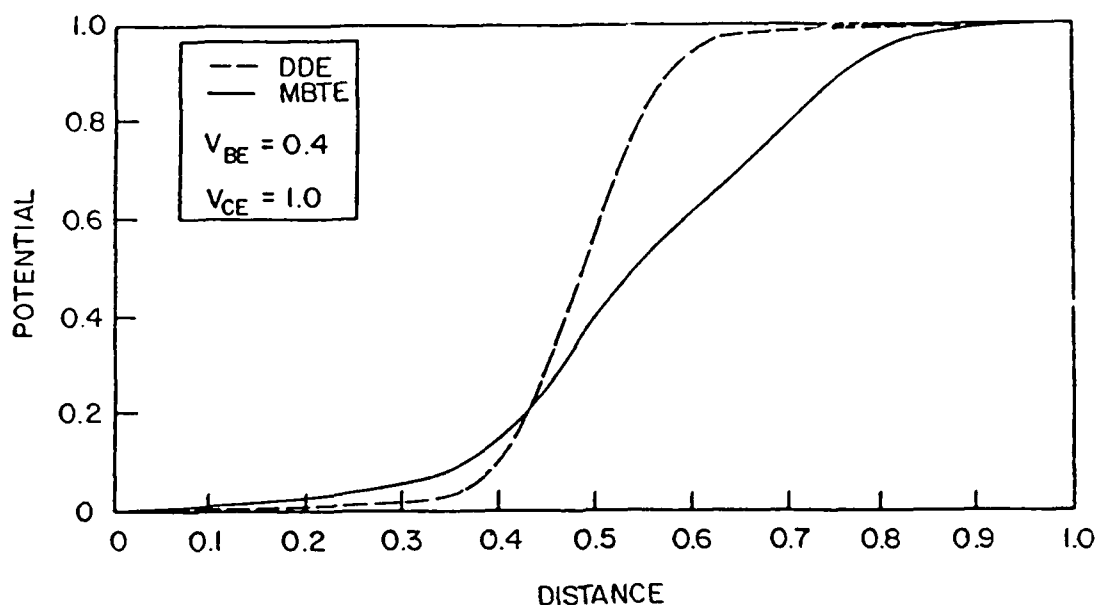exclusively Γ-valley transport. The implication is that if the Γ-valley
carrier velocity increases with increasing field, then the product of density
and cross sectional area normal to current flow must decrease to maintain
current continuity. Since the velocity increases faster than the area
decrease, the carrier density decreases.

At moderate bias levels typical FET calculations show a decreasing field as
the gate region is passed. This also occurs in the PBT. Now, as the cross
sectional area increases, the Γ-valley carriers exhibit a decrease in
velocity. It must be noted, however, that for the parameters of the
calculations the L-valley carriers make a negligible contribution to current.
Thus a decrease in carrier velocity results in a net increase in carrier
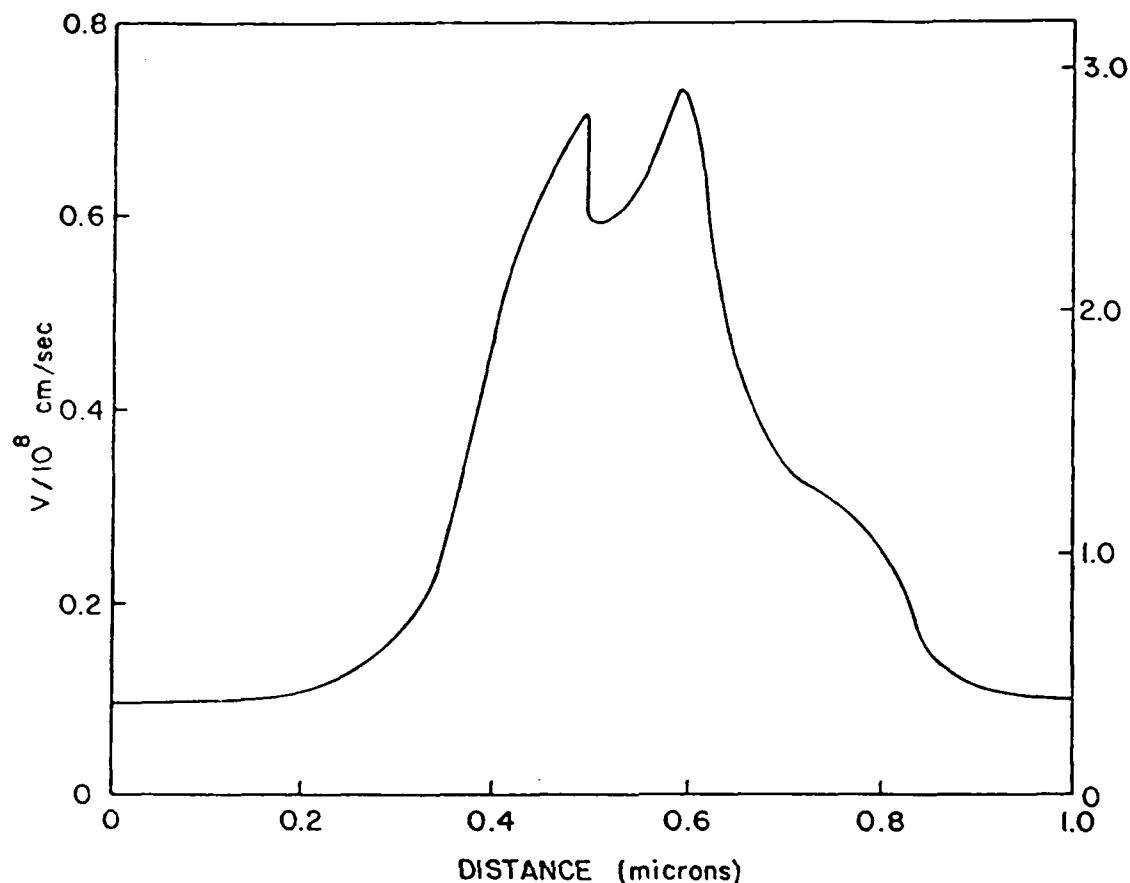
FIGURE 16: Γ-valley velocity along the center of the PBT channel [20].

concentration. However initially, the decrease in field is not accompanied by a corresponding temperature decrease (as experienced in the uniform field calculations). Thus, the high Γ-valley temperature results in transfer to the L-valley giving rise to the second minimum in the Γ-valley carrier density shown in Fig. 14. Further toward the drain, the field decreases. However, relaxation is incomplete and the field at the collector is not equal to the field at the emitter. Also note, the moment equation potential distribution gives rise to a slightly higher field upstream of the base, and a lower field, over a longer distance, downstream of the base compared to the drift and diffusion result. More significantly, the electron temperature at the collector exceeds that at the emitter. It is noted that as the field relaxes, the electrons transfer back to the central valley.

NONEQUILIBRIUM ELECTRON-HOLE TRANSPORT

Additional nonequilibrium studies were discussed at the beginning of this paper. This concerns nonequilibrium electron-hole transport, which for specificity was discussed for InGaAs. The details are considered below.

In this recent study, ensemble Monte Carlo studies were performed in which electrons were injected into p-type InGaAs [7]. In one case the acceptor doping was 10^{17} cm^{-3}, and in the second case 5×10^{17} cm^{-3}. The calculations were at 300°K and the ratio of the injected electrons to the majority holes was taken to be 1:5. (Note: The ensemble Monte Carlo avoids any assumptions on the magnitudes of the energy and momentum exchange in an electron-hole (e-h) scattering process and the evolution of the electron and hole distribution functions.) The electrons and holes were assumed to be in equilibrium with the lattice when the electric field is switched on; and the band model consisted of three nonparabolic valleys for the conduction band and two parabolic light- and heavy-hole bands. The role of the light holes was suppressed in this study. The model includes the elastic acoustic phonon, impurity scattering using Ridley's model, alloy scattering, deformation potential, intervalley and intravalley phonon scattering processes. The e-h and screened carrier-phonon scattering are calculated from the expressions given in [22], using a self-consistent screening model. Only one LO phonon mode in this calcu- lation. The electron transport parameters for $In_{0.53}Ga_{0.47}As$ are the same as those reported in [23]. However, for hole effective masses and deformation potential constants, constants appropriate to GaAs were used. The interaction between the L-valley electrons and the heavy holes was ignored.

The drift velocities of the electrons injected into the p-type InGaAs as a function of the applied electric fields are plotted in Figs. 4 and 17a, for doping levels of 5×10^{17} and 10^{17} cm^{-3}, respectively. The curves connecting the open circles in these figures correspond to situations where the interaction with the mobile holes is ignored and only the interaction with the ionized acceptor impurities is taken into account.
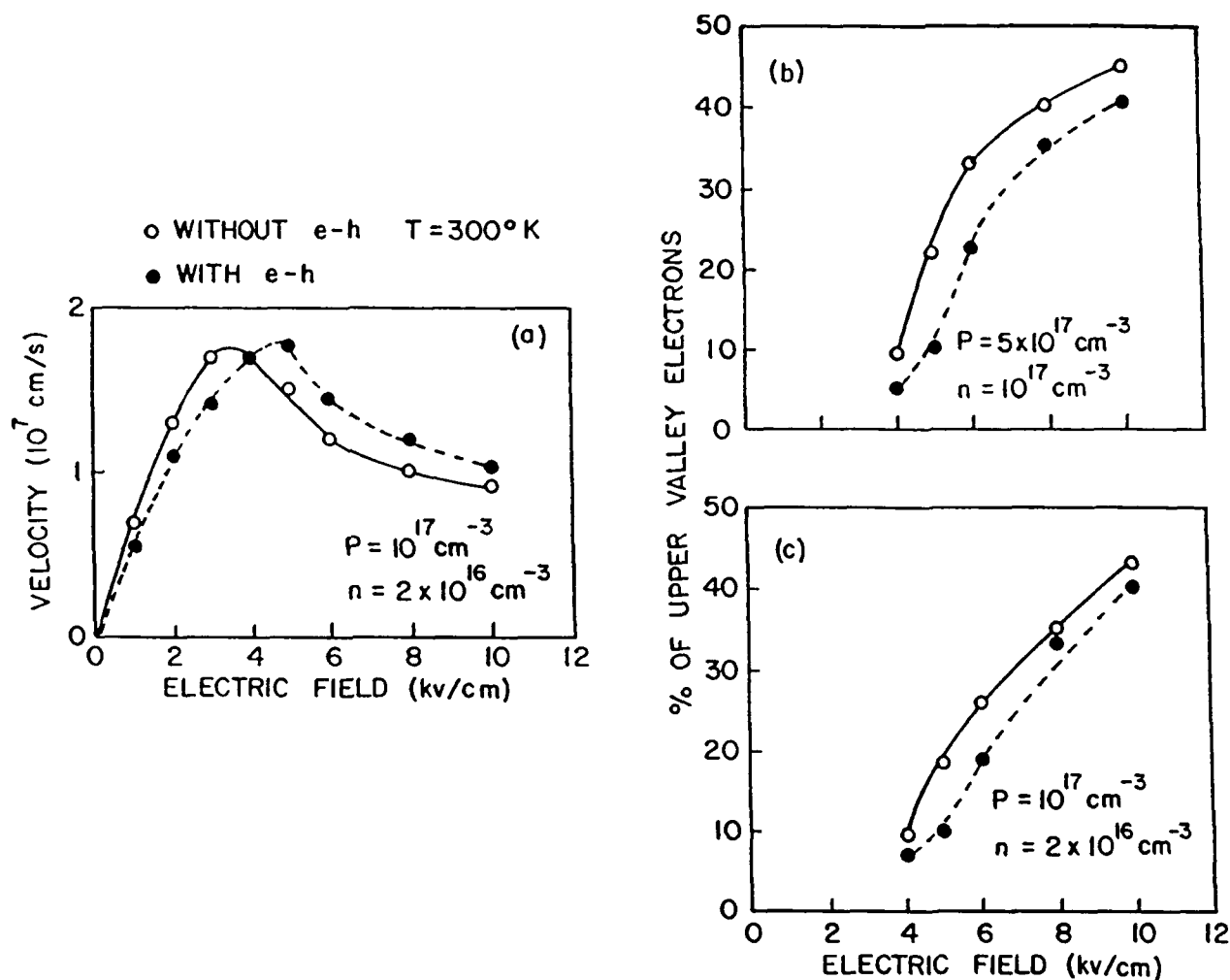
FIGURE 17: a) Electron drift velocity in p-type InGaAs.
            b) and c) Percentage of upper valley electons in p-type InGaAs [7].

When the interaction between the minority electrons and the mobile majority holes is taken into account, the electron velocities are lower for fields below 4 kV/cm compared to majority electrons as can be seen from these figures. At these low fields the electron transfer to the upper valleys is negligible as shown in Fig. 17b, and the energy loss through e-h interaction is not significant because the rate at which the electrons gain energy from the electric field is small as can be seen in Fig. 18. However, the e-h scattering which has the same angular distribution as the impurity scattering, has the same effect on the electron mobility as doubling the doping level. Consequently, the mobility of the electrons is reduced, leading to lower

velocities. As the electric field is increased, the fraction of electrons
with enough energy to transfer to the upper valleys increases.
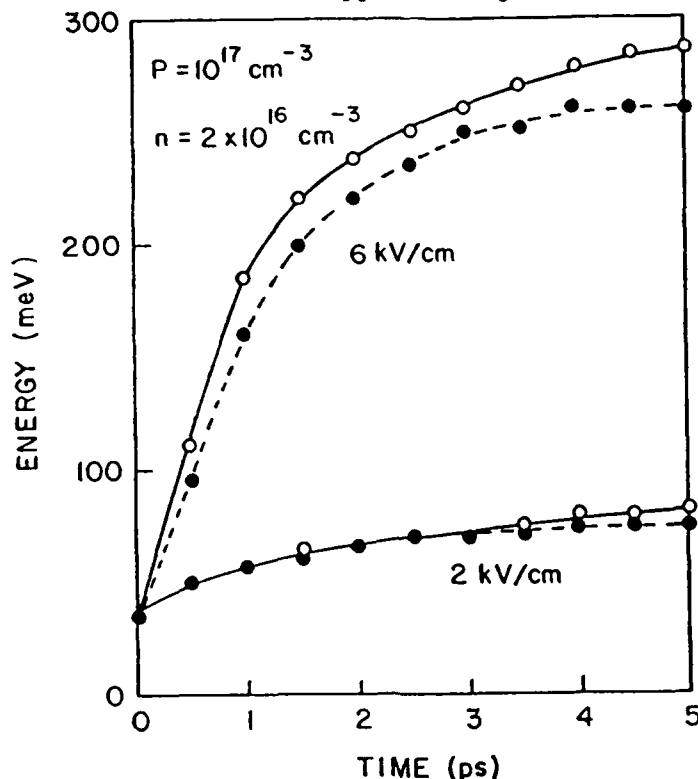
FIGURE 18: Electron energy in p-type InGaAs, as a function of field [27].

However, as the energy of the electrons increases, the rate at which they lose
energy through e-h scattering is increased [27]. This results in retaining
more electrons in the central valley when scattering of the electron by the
majority holes is taken into account as can be seen in Fig. 17, and results in
higher electron velocities above 4 kV/cm in the present situation. An
additional feature of the velocity field curve of the minority electrons is
that it converges to that of the majority electrons at higher electric
fields. This reflects the fact that at these high fields the rate at which
electrons gain energy from the electric field exceeds the rate at which
electrons lose energy to the heavy holes. Consequently, the population of
electrons in the upper valleys increases.

It is worthwhile to note that the experimental measurement [8] of minority
carrier velocity does not exhibit any negative differential resistance. The
origins of this are unclear. We do point out that the measurements are

performed for two-terminal systems and the highly nonuniform field distribution may prevent the appearance of NDR.

TRANSPORT IN ULTRA-SUBMICRON DEVICES

The entire discussion of transport has been predicated on a semi-classical picture. Certainly quantum effects take place on short-time scale where the Fermi-Golden rule breaks down and where spatial feature sizes are the order of tens of angstroms. As a rule of thumb, it is thought that quantum mechanical effects become prominent when the feature size is of the order of a thermal deBroglie wavelength or shorter, as shown below:
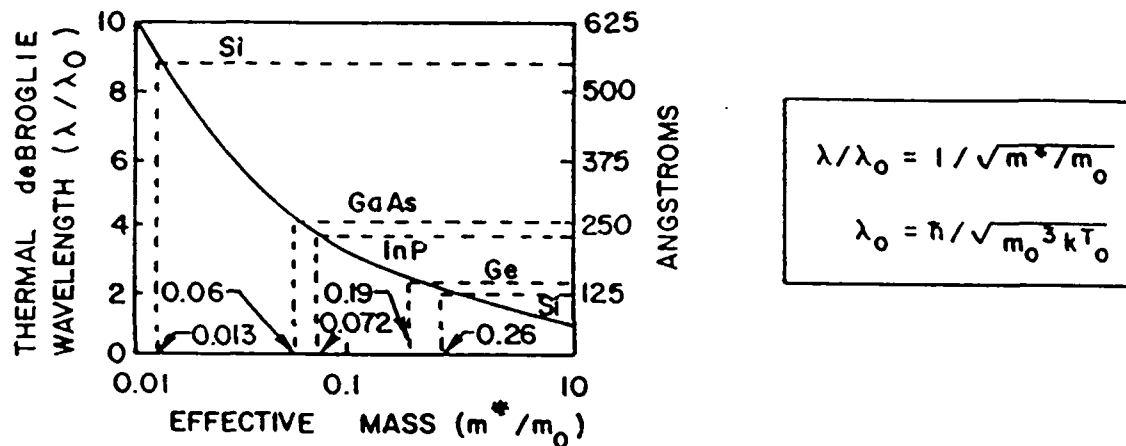
FIGURE 19: Thermal deBroglie wavelength vs effective mass [24].

The quantum transport formulation for devices is extremely rich and nes approaches are necessary. For example, it appears necessary to resort to solutions, e.g., to the density matrix or some equivalent form as the Wigner distribution function [25]. Moment equations are also applicable. For example, from the equation of motion of the density matrix, for a system including mobile carriers and scattering centers, the first three moment equations have the following form [26],

$$<<\dot{P}^{(0)}>> + \frac{1}{m} \frac{\partial}{\partial x} <<P^{(1)}>> - \frac{i}{\hbar} <<[H_s, P^{(0)}]>> \tag{12}$$

$$<<\dot{P}^{(1)}>> + \frac{1}{m}\frac{\partial}{\partial x}<<P^{(2)}>> = -\left(\frac{\partial V}{\partial x}\right)<<P^{(0)}>> + \frac{i}{\hbar}<<[H_s,P^{(0)}]>> \tag{13}$$

$$<<\dot{P}^{(2)}>> + \frac{1}{m}\frac{\partial}{\partial x}<<P^{(3)}>> = -2\left(\frac{\partial V}{\partial x}\right)<<P^{(1)}>> + \frac{i}{\hbar}<<(H_s,P^{(2)})>> \tag{14}$$

where the $<<\ >>$ denote quantum ensemble averages, and using Dirac notation, the operators of interest are of the form

$$P^{(0)} = |x_o><x_o| \tag{15}$$

$$P^{(1)} = \left(\frac{1}{2}\right)\left[P|x_o><x_o| + |x_o><x_o|P\right] \tag{16}$$

$$P^{(2)} = \left(\frac{1}{2}\right)^2\left[P|x_o><x_o| + 2P|x_o><x_o|P + |x_o><x_o|P^2\right] \tag{17}$$

$$P^{(3)} = \tag{18}$$

$$\left(\frac{1}{2}\right)^3\left[P^3|x_o><x_o| + 3P|x_o><x_o|P^2 + 3P^2|x_o><x_o|P + |x_o><x_o|P^3\right]$$

where p is the momentum operator. We note that the terms involving H_s incorporate dissipation. In a diagonal representation, the ensemble average of the first three operators breaks down into the following form.

$$<P^{(0)}> = \sum \rho_{ii}\ n_i(x_o) = n(x_o) \tag{19}$$

$$<P^{(1)}> = \sum \rho_{ii}\ n_i m v_i = n(x_o)m v_d \tag{20}$$

$$<P^{(2)}> = \sum \rho_{ii}m^2[(v_i-v_d) + v_d]^2 - \frac{\hbar^2}{4}\sum \rho_{ii}n_i\frac{\partial^2}{\partial x^2}\ln n_i$$

$$= \Omega_{xx} + m^2 n v_d^2 - \frac{\hbar^2}{4}\sum \rho_{ii}n_i\frac{\partial^2}{\partial x^2}\ln n_i \tag{21}$$

where ρ_{ij} is the diagonal element of the density matrix, and

$$\Omega_{xx} - \sum \rho_{ii} m (v_i - v_d)^2 \tag{22}$$

It is clear that with the exception of the third operator, which contains a term involving Plank's constant the equations are of a classical form. Thus, in the simplest approximation, there appears to be a close similarity between the classical moment equations and that obtained quantum mechanically. The difficult is of course in solving these equations.

There is however an interesting situation to consider: that in which $\rho_{ii} - 1/N$ the system. In this case the first two moment equations, including dissipation in momentum, reduces to

$$\frac{\partial n}{\partial t} + \text{divnv} - 0 \quad , \tag{23}$$

$$\frac{\partial nv}{\partial t} + \frac{\partial}{\partial x} nv_d^2 + \frac{nv}{\tau} - -\left(\frac{\partial V}{\partial x} + \frac{\partial Q}{\partial x}\right)\frac{n}{m} \tag{24}$$

where

$$Q - -\frac{\hbar^2}{2m} \frac{1}{\sqrt{n}} \frac{\partial^2 \sqrt{n}}{\partial x^2} \tag{25}$$

The quantity V represents an imposed barrier and the self-consistent energy associated with Poisson's equation. The potential Q [27] is density dependent and tends to become significant near strong barriers, where the curvature of $\sqrt{n}$ will either enhance or diminish the imposed barrier. Tunnelling, and resonance arises from Q.

In multiple dimensions these equations are subject to the constraint

$$\oint mv \cdot dx - nh \tag{26}$$

or in a gauge that includes a vector potential, the constraint

$$\oint \left(mv + \frac{e}{c}A \right) \cdot dx = nh \qquad (27)$$

Presently, device systems are being constructed which are influenced by the constraint of Eq. (27), often called the Aharonov-Bohm condition [28]. In particular, structures are bing constructed in which the path of an incident beam of electrons is split and then recombined. The split path lengths of the original beam are different, and under coherent reconstruction in which Eq. (27) is satisfied, conduction oscillations are anticipated. A structure originally proposed to deal with this is displayed in Fig. 20.
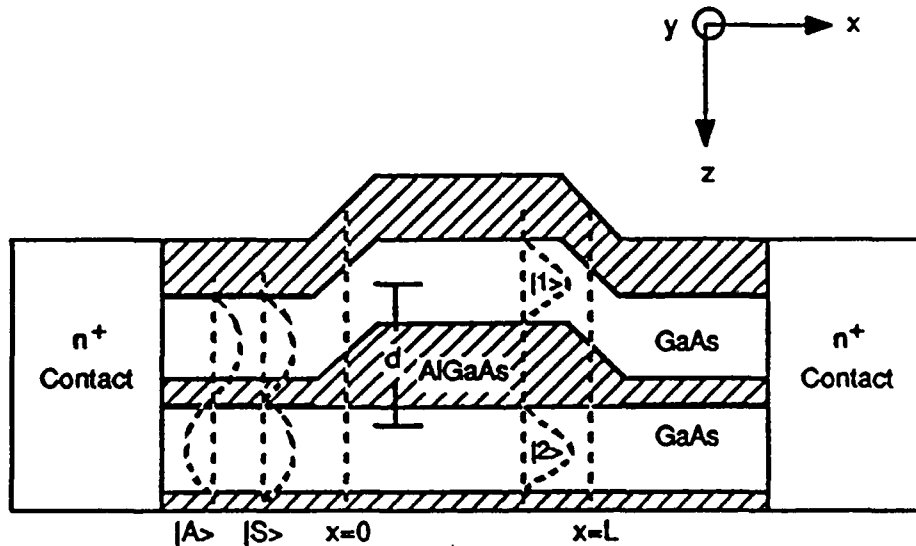

FIGURE 20: Configuration suitable for the Aharonov-Bohm constraint [29].

- 32 -

REFERENCES

1. J.B. Gunn, Solid State Commun. 1, 88 (1963).

2. J.G. Ruch, IEEE Trans. Electron Devices, ED-19, 652 (1972).

3. See, e.g. A.A. Ketterson, W.T. Masselink, J.G. Gedymin, J. Klem, C-K. Peng, W.F. Kopp, H. Morkic and K.R. Gleason, IEEE Trans. Electron Devices, ED-33, 564 (1986).

4. J.R. Chelikowsky and M.L. Cohen, Phys. Rev. B14 556 (1976).

5. H.L. Grubin, Proceedings of NATO Conference.

6. V.L. Dalal, A.B. Dreeben, A. Triano, J. Appl. Phys. 42, 2864 (1971).

7. M.A. Osman and H.L. Grubin, Appl. Phys. Letts.   (1987).

8. R.J. Degani, R.F. Leheny, R.E. Nahory and J.P. Heritage, Appl. Phys. Letts, 49, 1540 (1981).

9. Compiled from "Numerical Data and Functional Relationships in Science and Technology/ Landolt-Bornstein", Volume 22, Eds: O. Madelung and M. Schulz, Springer-Verlag, Berlin (1987).

10. From B.K. Ridley, "Quantum Processes in Semiconductors", Clarendon Press, Oxford (1982).

11. H.L. Grubin and J.P. Kreskovsky, Physica 134B+C, 67 (1985).

12. H.L. Grubin, D.K. Ferry and K.R. Gleason, Solid State Electronics, 23, 157 (1980).

13. See, e.g. P. N. Butcher, Repts. Prog. Physics, 30, 97 (1967).

14. See, e.g. K. Yamaguchi, S. Asai and H. Kodera, IEEE Trans Electron Dev. ED-23, 1283 (1976).

15. M. Meyyappan, J.P. Kreskovsky and H.L. Grubin (to be published).

16. M.P. Shaw, P.R. Solomon and H.L. Grubin, IBM J. Res. and Devel. 13, 587 (1969).

17. See, e.g. E.M. Azoff, Solid State Electronics, 30, 913 (1987).

18. See, e.g., R.K. Froelich Avionics Laboratory Report AFWAL-TR-82-1107.

19. C.O. Bozler and and G.D. Alley, IEEE Trans. Electron Dev. ED-27, 6 (1980).

20. J.P. Kreskovsky, M. Meyyappan and H.L. Grubin, Proceeding of NUMOS I, Ed. J.J.H. Miller Boole Press (1987).

21. R.F. Cook and J. Frey, COMPEL 1,2 (1982).

22. M.A. Osman and D.K. Ferry. J. Appl. Phys. 61, 5330 (1987).

23. S.R. Ahmed, B.R. Nag and M.D. Roy, Solid State Electronics, 28, 1193 (1985).

24. G.J. Iafrate, H.L. Grubin and D.K. Ferry, J. de Physique, 42-C7, C7-307 (1981).

25. E. Wigner, Phys. Rev. 40, 749 (1932).

26. H.L. Grubin (to be published).

27. See, e.g. C. Philippidis, D. Bohm and R.D. Kaye, Il Nuovo Comento, 71B, 75 (1982).

28. Y. Aharonov and D. Bohm, Phys. Rev., 115, 485 (1985).

29. S. Datta, M.R. Mellcoh, S. Bandyopadyay, and M.S. Lundstrom, Appl. Phys. Letts. 48, 487 (1986).

QUANTUM MOMENT BALANCE EQUATIONS AND RESONANT TUNNELLING STRUCTURES*

H.L.Grubin and J.P.Kreskovsky
Scientific Research Associates, Inc., P.O.Box 1058
Glastonbury, Connecticut, 06033-6058

Abstract

This study describes the evolution and implementation of a set of quantum balance equations for examining transport im mescoscopic structures.

Key Words

Wigner functions, quantum potentials, quantum balance equations.

Introduction

This study describes the evolution and implementation of a set of quantum balance equations for examining transport im mescoscopic structures. The study is motivated by a perceived need for an intuitively accessible set of multi-dimensional quantum transport equations, that permit the self-consistent calculation of particle current and current density. The goal is the development of a set of quantum hydrodynamic equations that reduce to the single particle equations [1] for a pure state, and the classical hydrodynamic equations [2] as $\hbar \rightarrow 0$. As discussed below, these goals have been partially met.

Pure State and Classical Moment Equations

The hydrodynamic equatons for a pure state, for single particle transport, spatial variations in one dimension, and a classical potential $U(x,t)$, are, with $\psi(x,t) = \sqrt{\rho} \exp[iS(x,t)/\hbar]$, and $p_d = \partial S/\partial x$:

$$\partial_t \rho + \partial_x(p_d \rho/m) = 0 \tag{1}$$

$$\partial_t(\rho p_d) + \partial_x(\rho p_d^2/m) + \rho \partial_x(U+Q) = 0 \tag{2}$$

$$Q = -(\hbar^2/2m\sqrt{\rho})\partial_x^2\sqrt{\rho} \tag{3}$$

In the above, ρ and p_d/m represent probability density and probability current flux; quantum mechanics is represented by the quantum potential, Q [3]. (Note: because of the one dimensional nature of the transport, 'Bohr-Sommerfeld' constraints are automatically satisfied [3].) The classical moment equations for single band transport and spatially independent effective mass are [2]:

$$\partial_t \rho + \partial_x(p_d \rho/m) = 0 \tag{4}$$

$$\partial_t(\rho p_d) + \partial_x(\rho p_d^2/m) + \rho \partial_x U + \partial_x(\rho kT) = \partial \rho p_{d,coll} \tag{5}$$

$$\partial_t W + \partial_x(p_d W/m) + \partial_x(p_d \rho kT/m) + (\rho p_d/m)\partial_x U = \partial W_{,coll} \tag{6}$$

$$W = 3\rho kT/2 + \rho p_d^2/2m \tag{7}$$

It is worthwhile emphasizing that the above equations involve three dimensional momentum space integration, with spatial variations in only one direction, and that density and momentum now represent particle density. The derivative notation in the above equations is $\partial_x = \partial/\partial_x$, etc.

Structure of the Quantum Mechanical Equations

If the quantum transport equations for a pure state are given by equations (1) through (3), and the classical equations are given by (4) through (7), at the very least it may be anticipated that quantum contributions will arise by replacing the classical potential $U(x,t)$ by $U(x,t) + Q(x,t)$. How good is this statement? To examine this we turn to an approximate non-equilibrium Wigner function, discussed in [4].

The non-equilibrium distribution function is constructed [4] from the zero current equilibrium distribution function obtained by Wigner [5], and discussed more recently by Ancona and Iafrate [6]:

$$f_w = \exp{-\beta[p^2/2m + U]}\{1-2\alpha(\partial_x^2 U-\beta(\partial_x U)^2/2)/3-\alpha(1-\beta p_x^2/m)\partial_x^2 U/3\} \tag{8}$$

In equation (8), $\beta = 1/kT$, $\alpha = \hbar^2\beta^2/2m$, and $p^2 = p_x^2 + p_y^2 + p_z^2$.

The construction of the nonequilibrium distribution function involves replacing the potential and its derivatives in equation (8) by corresponding density expressions. The carrier density is

$$\rho = 2(1/2\pi\hbar)^3 \int f_w(x,p)d^3p = N\exp{-\beta U}[1-2\alpha(\partial_x^2 U-\beta(\partial_x U^2)/2)/3] \tag{9}$$

where $N = 2(m/2\pi\beta\hbar^2)^{3/2}$. After demonstrating that $\partial_x U = -(\partial_x \rho)/(\beta\rho) + O(\alpha)$, and $\partial_x^2 U = -\partial_x\{(\partial_x\rho)/\rho)\}/\beta + O(\alpha)$, it is a direct matter to show:

$$f_0 = (\rho/N)\exp[-\beta p^2/2m][1+(\alpha/3\beta)(1-\beta p_x^2/m)(\partial_x\{(\partial_x\rho)/\rho\})+O(\alpha^2)] \tag{10}$$

Note: when equation (9) is substituted into equation (4), with U representing the equilibrium potential, equation (8), to order α, is retrieved.

To see what equation (10) offers, consider the steady state small signal Wigner function within a relaxation time approximation, and to second order in $\hbar$ [6],

$$f_w = f_0 - \tau_w[(p_x/m)\partial_x f_0 - (\partial_x U)\partial_p f_0 + (\hbar^2/24)(\partial_x^3 U)\partial_p^3 f_0] \tag{11}$$

Inserting equation (10) into (11), the following key results emerge: $\int f_w d^3 p = \int f_0 d^3 p$, and for $j_x = -(2e/(2\pi\hbar)^3)\int f_w(p_x/m)d^3 p$:

$$j_x = \mu\{\rho\partial_x(U+Q/3)+\partial_x(kT\rho)\} \tag{12}$$

which was first obtained by Ancona and Iafrate [6]. Here, $\mu = e\tau/m$. While this result is consistent with the general philosophy of the introductory paragraph of this section the factor of 3 on the quantum potential needs to be explored. For a displaced version of equation (10), the factor of 3 is retained for the moments of the Wigner-Boltzmann (WB) equation, as considered next.

Moments of the Wigner-Boltzman Equation

The quantum moment equations (see also [7]) have been obtained for the WB equation with quantum contributions to order $\hbar^2$, and for a displaced distribution function in which p in equation (10) is replaced by $p-p_d$. The WB equation of motion is:

$$\partial_t f_w + (p_x/m)\partial_x f_w - (\partial_x U)\partial_p f_w + (\hbar^2/24)(\partial_x^3 U)\partial_p^3 f_w] = f_{w,coll} \tag{13}$$

and the first three moment equaton corresponding to that of equations (4) through (7) are:

$$\partial_t\rho + \partial_x(p_d\rho/m) = 0 \tag{14}$$

$$\partial_t(\rho p_d)+\partial_x(\rho p_d^2/m)+\rho\partial_x(U+Q/3)+\partial_x(\rho kT) = \partial\rho p_{d,coll} \tag{15}$$

$$\partial_t W + \partial_x(p_d W/m)+\partial_x(p_d\rho kT/m)+(\rho p_d/m)\partial_x(U+Q/3)$$
$$-(\rho\hbar^2/12m)\partial_x\{(\partial_x\rho)/\rho\}\partial_x(p_d/m) = \partial W_{,coll} \tag{16}$$

$$W = 3\rho kT/2 + \rho p_d^2/2m-(\rho\hbar^2/24m)\partial_x\{(\partial_x\rho)/\rho\} \tag{17}$$

Equation (17) and equation (9) for density has strict quantum mechanical meaning (see also the discussion following equation (25) in [5]). Before attending to the above results it is important to establish a confidence level in the quantum balance equatins. To this end the general moment equation formulation, including disspation, of Stroscio [8] is recalled. Stroscio's results while specific to a phase space that includes one space dimension and one momentum direction, overlap those of this study, as demonstrated below. In this case, borrowing the notation of [8] within the framework of the displaced Wigner function used herein, it is straightforward to first show that:

$$\rho <(p-p_d)^2> = m\rho[1-(2\alpha/3\beta)(\partial_x\{(\partial_x\rho)/\rho\}/\beta] \tag{18}$$

$$<\rho(p-p_d)^3> = 0 \tag{19}$$

from which reference [8] equations (10a), and (10b) when combined with (10a) yield equations (14) and (15) of this study, where the collision integrals are treated generically. The energy balance equation is treated similarly. Here, reference [7] equation (10b), when multiplied by $\rho p_d/m$ is added to (10c), which is multiplied by 1/2m; the continuity equation is included in this procedure. The result of this manipulation is a one dimensional phase space version of the energy balance equation of this study, W is replaced by:

$$W' = \rho kT/2 + \rho p_d^2/2m-(\rho\hbar^2/24m)\partial_x\{(\partial_x\rho)/\rho\} \tag{20}$$

From the point of view of device modeling, it is pointed out a quantum corrected quasi-Fermi energy can be defined. Writing $E = U + Q/3 + k\mathcal{U}n[\rho/\rho_r]$, where ρ_r is a reference density, the current density in equation (12) can be written as $j = \mu\rho\nabla E$ [6], and equations (15) and (16) can be reexpressed as:

$$\partial_t(\rho p_d)+\partial_x(\rho p_d^2/m)+\rho\partial_x E = \partial\rho p_{d,coll} \tag{21}$$

$$\partial_t W + \partial_x(p_d W/m) + \partial_x(p_d\rho kT/m) + (\rho p_d/m)\partial_x E$$
$$-(\rho\hbar^2/12m)\partial_x\{(\partial_x\rho)/\rho\}\partial_x(p_d/m) = \partial W_{,coll} \tag{22}$$

In the above form it appears that the dynamics of the transport are governed by an energy E. However, E is introduced as a transformation of variables.

Anticipated Solutions of the Quantum Balance Equations

We focus attention on the single particle pure state equations, whHere of a zero time derivative of the momentum balance equation implies that $S(x,t) = S_1(x) + S_2(t)$, and $\rho(x,t) = \rho(x)$. For zero time derivatives, one space integration yields (ignoring spatial derivatives of the effective mass) energy conservation, $p_d^2/2m + (U+Q) = E_o$, and constant probability current, $\rho p_d/m = J$, where E_o and J are integration constants. Using the definition of the quantum potential, energy conservation can be rewritten as:

$$\partial_x^2 J\rho + (2m/\hbar^2)[E_o - U - (mJ^2/2\rho^2)]J\rho = 0 \qquad (23)$$

For bound states, $J = 0$, and equation (23) is an eigenvalue problem, one that in the case of a resonant tunnel structure leads to quasi-bound states. Further under zero current conditions, with E_o representing the eigenvalues, $Q = E_o - U$, and the values of Q are spatially dependent and, in some cases, are approximately equal to the bound states. This result will be prominantly displayed in the discussion below.

The single particle Schrodinger picture is limited, in that being dissipationless it does not permit a direct transition transition to a multiparticle problem when contacts are considered. For example, in the case of multiparticle transport with electrons moving ballistically within the N^- region of an $N^+N^-N^+$ structure, the mean carrier energy and velocity increase from the cathode to the anode. Conservation of multiparticle current requires that increases in velocity are accompanied by decreases in particle density. Thus in the absence of dissipation there will necessarily be charge depletion at the downstream anode, unless dissipation is present in the interior of the device. If the assumption is made, that the physical contact are boundaries where the numbers of carriers at the cathode and anode are equal, then scattering within the interior of the structure is conceptually necessary. For the hydrodynamic formulation of the single particle Schrodinger's equation, there is no meaning to introducing N^+ cathode and anode regions, since we are dealing with a single particle.

In that dissipation is an essential feature of transport in devices, the quantum balance equations represented by equations (13) through (17) form a staring point for the simulations to be discussed below. To date, our simulations include the first two moment equations, and Poisson's equation. These have been solved for a spatially dependent effect mass, and for Fermi statistics. Here, since we have neither generated a set of WB moment equations for a spatially dependent effective mass nor have we obtained a displaced Wigner function that satisfied Fermi statistics, we have instead patched on these contributions. Further, we have treated the factor of '3' associated with equations (14) and (15) as an adjustable parameter that reflects the statistical distribution used as a basis for the calculation, as discussed in [6], and have replaced it by unity. In this case with $v = p_d/m$, the continuity equation is unchanged, while the momentum balance equation reads: :

$$\partial_t \rho v + \partial_x(\rho v^2) + (\rho/m)\{[\partial_x(U+Q)] + [(\rho v^2/2) - (NkTF_{3/2}/\rho)]\partial_x \ell \, nm\} + (2/3)\partial_x NkTF_{3/2} + \rho v\Gamma = 0 \qquad (24)$$

where $F_\ell(x_f) = (2/J\,\pi)\int [x^\ell/(1+\exp(x - x_f))]dx$, where the integration range is $o < x < \infty$. 'x_f' is defined implicitly as a change of variables through the relation $F_{1/2} = \rho/N$; where $x_f \approx \ell \, n(\rho/N) + \rho/(N\sqrt{8})$, for $x_f < 4.4426$, and $x_f = (9\pi/16)^{1/2}(\rho/N)^{2/3}$, for $x_f > 4.4426$. Using the identities associated with the Fermi integral, namely $F_{3/2} = (2/3)F_{1/2}$, and introducing the term $E = U + Q + kTx_f$, which is a generalization of the variable transofrmation discussed above the momentum balance equation is:

$$\partial_t \rho v + \partial_x(\rho v^2) + (\rho/m)\{[\partial_x E] + [(\rho v^2/2) - (NkTF_{3/2}/\rho)]\partial_x \ell \, nm\} + \rho v\Gamma = 0 \qquad (25)$$

Equation (25) is coupled to the equation of continuity and Poisson's equation, with U(x) representing the conduction band energy. The heterostructure is represented by the Anderson rule: $U = \Sigma - \chi(x)$, where $\chi(x)$ is a position dependent electron affinity. Σ is obtained from Poisson's equation: $\nabla\epsilon\nabla\Sigma = e^2[\rho - \rho_o]$, where $\epsilon(x)$ is a position dependent permitivity, and ρ_o is a position dependent doping level. For conduction band variations between GaAs and $Al_xGa_{1-x}As$, the following relationships were used: $m = 0.067 + 0.083x$, $\Delta E_c = 0.697x$.

Calculations

The calculations discussed below are for the structure shown in figure 1, with resonant tunnelling barriers located symmetrically at the center of the structure. The structure and dimensions of the barriers are displayed in figure 2, which shows the current voltage characteristics of this device at 77K. There is a weak region of negative conductance. The conduction band profile at different bias levels, figure 3, shows the expected tilt as the bias is increased. The distribution of energy is such that at 0.1v, approximatley 20% of the voltage drop falls across the upstream accumulation layer, 30% within the confines of the barrier, and 50% across downstream from the second barrier. The charge distribution, figure 4, shows a region of charge accumulation upstream of the barrier that increases with increasing bias, as does the charge in the well. While different boundary conditions have not been studied these results should be extremely sensitive to the boundary conditions at the cathode; as should the effects of incorporating the energy balance equation. It is not clear that including the latter will reproduce the charge depletion at resonance seen by several other studies.

In all our simulations, we have noticed the formation of a depletion layer downstream of the second barrier once we pass the valley current of the I-V characteristic. This depletion layer is a specific single particle quantum effect. The depletion layer keeps on extending for biases greater than the voltage at the valley of the I-V curve until the depletion layer touches the heavily doped region ($2 10^{18} cm^{-3}$). Then, the electron density downstream of the second barrier gradually increases and the depletion region disappears.

The quantum potential is displayed in figure 5. If we concentrate on its value in the well, the most dramatic point to note is that as the bias is raised the value of the quantum potential tends to cluster around a narrow range, increasing in magnitude from the upstream barrier to the downstream barrier. Within the barrier the

values of Q cluster around the quasi-bound state value associated with single particles. This result tends to emphasize that the approach to resonance is governed by single particle contributions. We note that Q is obtained from and affects ρ. Here, at low values of bias ρ in the first barrier is small, this is concomitant with large values of Q in the same region. The large value of Q also prevents the carriers from moving through the second barrier and forces them to pile up in the well. For larger values of bias the strength of Q decreases in the second barrier and charge density begins to buildup.

High values of carrier density in the well are consistent with low values of carrier velocity, and vice versa. It is found that at low values of bias the average time spent in the well, defined as L/<v>, where L is the width of the well, and <v> is the mean velocity in the well, varies at low values of bias from 10fs, to approximately 500fs at bias levels in the region of negative differential resistance. Indeed, the calculations suggest that it is the significantly reduced velocity at resonance that is responsible for NDR. The low average velocity in the well in contrast to the very high average velocity in the barrier reflects the fact that the electron spends more time in the well than in the barrier. Correspondingly, the charge density in the well is more prominent than in the barriers. The average velocity is always large in the depletion region after the second barrier. Indeed, velocities in the second barrier reached their saturated drift value; a result consistent with the proposal by Luryi [9]. We note that while the well density increases with bias, an estimate of the integrated charge indicates that it decreases the structure decreases in the NDR region, with most of the decrease in local charge occuring in the downstream barrier and the region between the downstream barrier and the collector contact.

Conclusions

The calculations display several very distinct features: (1) The charge from the cathode tunnels through the first barrier into the well, and the amount of charge in the well increases with increasing bias. (2) Resonance and NDR are dominantly single particle effects and are accompanied by a dramatic decrease in velocity in the well. (3) Excessive increases in velocity in the second barrier must be accompanied by either elastic and/or inelastic scattering to prevent the mean velocity from reaching unrealistically high values. (4) The current voltage relationships exhibit peak-to-valley ratios which are smaller than the experimental values. However, for a device with a cross-section of 25μm x 25μm, a typical experimental cross-section, the peak-current calculated numerically is of the order of 0.2-20 mA. This is the typical range of peak current of various RTDs studied experimentally. The peak-to-valley ratios are typically a factor 2 to 4 lower than the experimental findings. This probably results from the inclusion of a relaxation term in the momentum balance equation. It has been shown experimentally that the presence of scattering in typical RTD (voluntary doped) could substantially reduced their peak-to-valley ratios. (5) The peak current occurs at a rather small bias and its location is strongly dependent on the actual doping profile outside the resonant tunneling structure itself. The results are strongly dependent upon the details of the quantum distribution function, whose form was assumed rather than calculated.

Summarizing, even though incomplete, this study is the first implementation of the first two quantum balance equations. These equations lead to occurence of a NDR region in typical resonant tunneling structures. Further work is needed to include the energy-balance equation, together with a more sophisticated treatment of scattering (both elastic and inelastic). When included, those refinements will allow us to distinguish between the possibilities of coherent and/or sequential tunneling in resonant tunneling structures, and be used as a powerful tool to design typical RTDs.

Acknowledgements

References

[1] See e.g., 'Quantum Mechanics' A. Messiah, North Holland, Amsterdam (1961).
[2] H. L. Grubin and J. P. Kreskovsky, VLSI Electronics, 10, 237 Academic Press, NY (1985)
[3] C. Philippidis, D. Bohm and R. D. Kaye, IL Nuovo Cimento, 71B, 75 (1982)
[4] M. G. Ancona, G. J. Iafrate and H. L. Grubin. to be published.
[5] E. P. Wigner, Phys. Rev. 40, 749 (1932).
[6] M. G. Ancona and G. J. Iafrate, Phys. Rev. B15 39, 9536 (1989).
[7] M. G. Ancona and H. F. Tiersten, Naval Research Laboratory Memorandum Ret 6444, March 1989.
[8] M. A. Stroscio, Superlattices and Microstructures 2, 83 (1986).
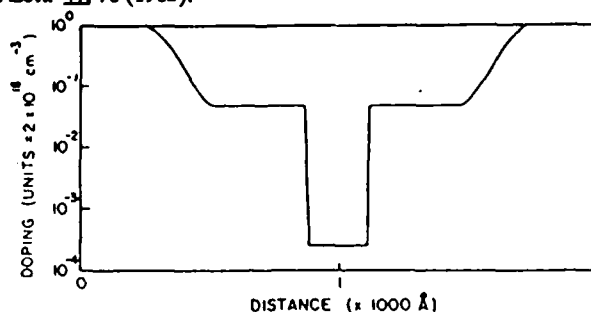[9] S. Luryi, Appl. Phys. Lett. 47, 70 (1985).

Figure 1. Doping profile in the numerical simulations. The RTD is at the center of the device
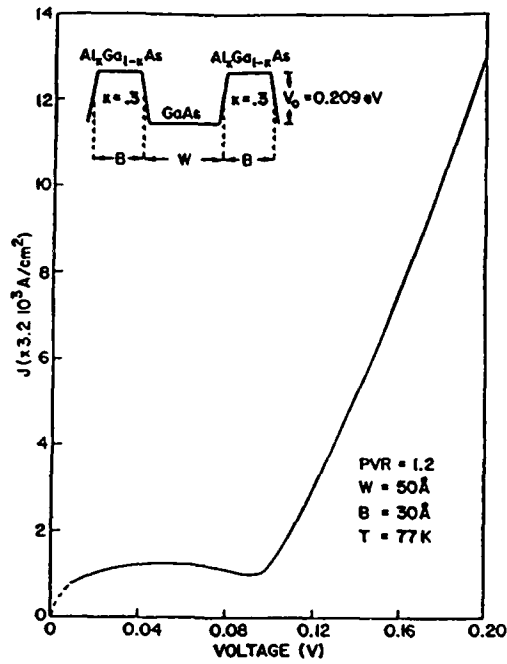
Figure 2. Current-Voltage Characteristic for the RTD
with the parameters shown in the inset. The
dashed line is the expected shape of the I-V
curve at low value of bias.
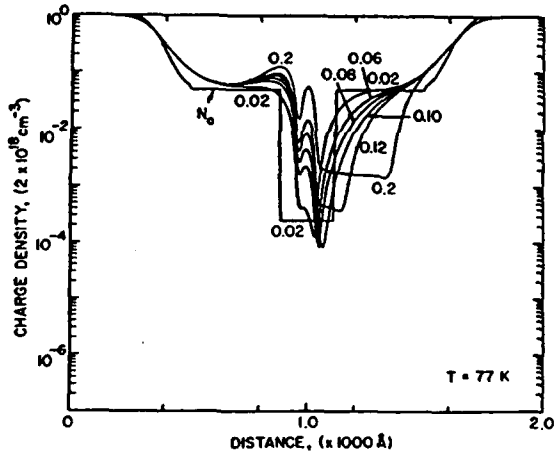
Figure 4. Charge density as a function of bias. N_0 is
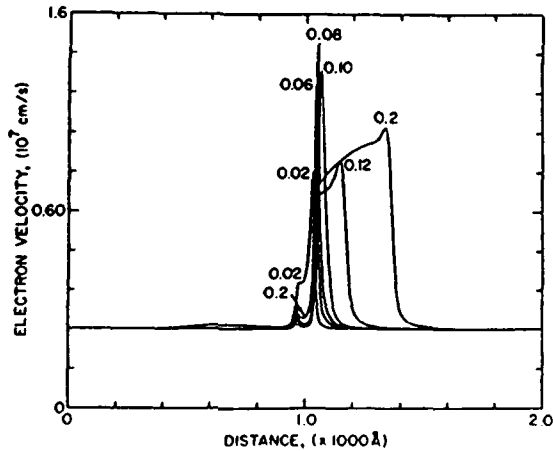the doping profile.
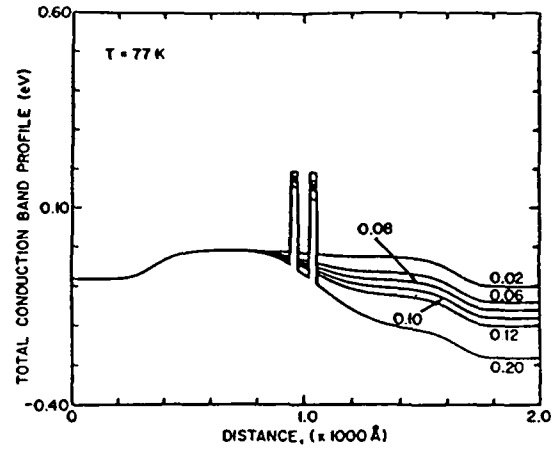

Figure 6. Average velocity as a function of bias.

Figure 3. Conduction band energy profile as a function
of bias. The different curves are labeled with
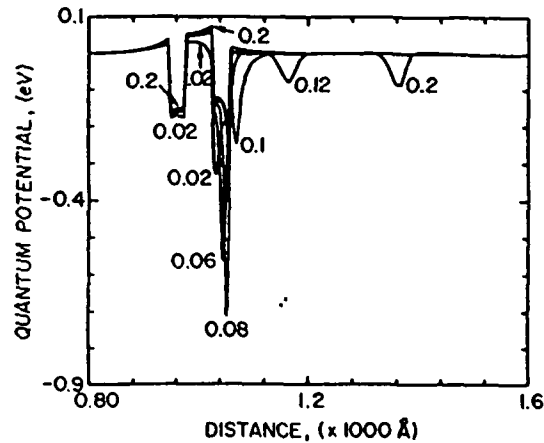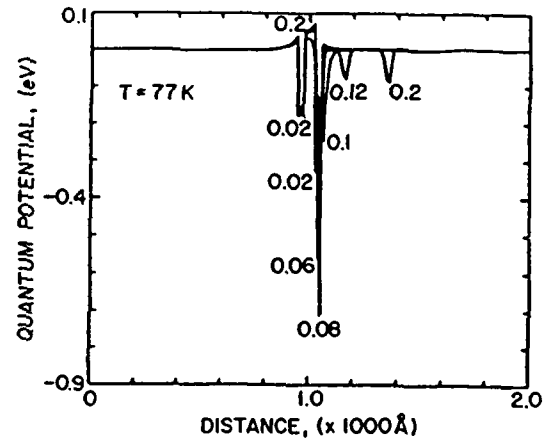the value of the voltage at which they are
calculated.

Figure 5. Quantum potential as a function of bias. The
lower figure is a closer look at the quantum
potential variation shown in the upper figure.

Appendix C

ELECTRON DIFFRACTION THROUGH
AN APERTURE IN A POTENTIAL WALL

M. Cahay, J. P. Kreskovsky and H. L. Grubin

Scientific Research Associates, Inc., P. O. Box 1058
Glastonbury, CT 06033

Abstract

We study the diffraction of a two-dimensional Gaussian wavepacket through a rectangular aperture in a finite potential wall (one slit experiment). For wavepacket with incident wavevector k_0 satisfying the diffraction condition, $k_0 = 2\pi/w$, w being the slit width, the near field (Fresnel-like) diffraction pattern behind the slit can be clearly seen for small time duration (< 0.2ps). At later time steps, the diffracted beam is fragmented into lobes (perpendicular to the direction of incidence of the wavepacket) as a result of the multiple reflections of the wavepacket inside the slit (assumed to be of finite thickness). At later time, no far-field Fraunhoffer diffraction pattern is observed in our numerical simulations.

Keywords

Electron diffraction; Schroedinger equation; ballistic transport; split gate.

Introduction

With the advent of sophisticated growth techniques such as Molecular Beam Epitaxy and Metal Organic Chemical Vapor Deposition, there has been an increased theoretical interest in various quantum mechanical tunneling problems including: (1) the resonant tunneling of electrons through double barrier heterostructures, a problem of primary importance in asserting the switching time of resonant tunneling device (RTD) (Hauge and Stovneng, 1989; Goldberg and coworkers, 1967), (2) the electron propagation through narrow ballistic constriction (defined by a split-gate) in the two-dimensional electron gas of a GaAs-Al$_x$Ga$_{1-x}$As heterostructure. The latter followed the recent experimental discovery (Van Wees and coworkers, 1988) that the conductance of such constrictions increases in a sequence of steps of height $2e^2/h$ (at sufficiently low temperature). More recently, the possibility of using narrow split-gate for transistor applications has been suggested by Kriman and coworkers (1988) in their newly proposed QUADFET.

Both RTD's and QUADFET's have potential high-speed device applications within the terahertz regime (Hauge and Stovneng, 1989; Bandyopadhyay and coworkers, 1989). On one hand, the fast switching response of RTD's has been widely investigated using purely one dimensional analysis of wavepacket propagation through double barrier heterostructures (Hauge and Stovneng, 1989; Goldberg and coworkers, 1967). In practice however, electrons are injected from 3D contacts and tunneling through the quantum well of the RTD is characterized by two-dimensional dynamics. On the other hand, in view of their potentiality for high-speed applications (such as in the QUADFET), there is now a call for a transient analysis of quantum transport through narrow ballistic constrictions to supplement the two dimensional steady-state analysis completed recently by several authors (Szafer and Stone, 1989; Kirczenow, 1989 and references therein).

Hereafter, we limit our numerical simulations to the diffraction of a gaussian wavepacket through a narrow aperture in a potential wall of finite height. The thickness d and length w of the slit are assumed to be 100Å and 200Å respectively, which corresponds to an aspect ratio a = d/w equal to 0.5. The choice of such a narrow slit was imposed to make the problem tractable numerically without having to use an excessive number of grid points. However, the time evolution of the diffracted beam shows interesting features which could still be present in the electron diffraction through a more realistic split gate realizable with present-day technology (Thornton and coworkers, 1986; Zheng and coworkers, 1986). The use of electron diffraction through a split-gate was recently proposed by Kriman and coworkers (1988) for interesting device applications.

The Numerical Approach

Recently, Ancilatto and coworkers (1989) have developed a method to solve the multi-dimensional Schroedinger equation based on the Chebychev polynomial expansion of the time evolution operator. Each term in this expansion is calculated using FFT computations. The subsequent effort to calculate the quantum-mechanical wavefunction $\psi(t)$ scales roughly as $\sim MN^d \ln N$, where M is the number of terms in the Chebychev expansion, d is the number of space dimensions and N^d is the total number of grid points. One drawback of this approach is that it requires the use of uniformly spaced grid points. Another drawback of the Chebychev expansion of the time-evolution operator is that, being a truncated series expansion, it does not conserve the unitarity of the time-evolution operator unless the truncated series converges. As a consequence, energy and norms of the wavefunction need not be conserved at any time step.

In this paper, we used an alternative approach to solve the time-dependent Schroedinger equation based upon a finite-difference solution procedure to solve a set of coupled equations governing the real and imaginary components of the wavefunction. The use of Crank-Nicolson differencing scheme insures conservation of the norm of the wavefunction at all times. The resulting coupled difference equations are then solved using a block alternating direction implicit (ADI) technique following the scalar ADI development of Douglas and Gunn (1964). Recently, a similar algorithm has been used by Barker (1989) to study the wavepacket propagation through Aharanov-Bohm rings. The technique can be used with non-uniform grid spacings, allows for an explicit time-dependence of the potential energy profile and can readily be extended to include the presence of a spatially varying effective-mass, of an external magnetic field and to three dimensional configurations (Cahay and coworkers, 1989).

Numerical Examples

As an illustrative example, we consider the diffraction of a two-dimensional Gaussian wavepacket through a narrow constriction such as the one shown in the upper left frame of Fig. 1. The subsequent frames show the time-evolution of the wavepacket impinging on a 100 Å wide potential barrier of height 0.06eV containing a 200 Å wide slit where the potential is assumed to be zero. The slit is disposed symmetrically with respect to the center of the simulation domain, a square of dimensions 3,000Å x 3,000Å. Our numerical simulations were performed using a non-uniform grid spacing with 69 and 79 grid points in the x and y directions respectively. In an actual split-gate, the slit can actually be varied from a few hundred to a few thousand angstroms while sweeping the gate voltage. The potential in the 2D electron gas is also different from the sharp potential wall assumed in the present work. We will comment on this point in our conclusion section. The initial wavepacket is assumed to be

$$\psi(x,y,t=0) = \left[1/\sqrt{\pi\sigma^2}\right] \exp[ik_0 x] \exp[-((x-x_0)^2+(y-y_0)^2)/2\sigma^2] \tag{1}$$

Where (x_0,y_0) are the coordinates of the wavepacket center and σ is equal to 100Å. Finally, the electron wavevector k_0 is 0.0094 Å^{-1} (with this value of k_0, a free electron ($m^* =0.067m_0$) travels a length of 1,000Å in 0.6 ps). The average kinetic energy of an electron in the state (1) is about 10 meV for the values of the parameters listed above and is therefore about 1/6th of the barrier height. A Fermi energy of about 10 meV corresponds to a typical 2D electron gas sheet density of ~ 3×10^{11} cm^{-2}. In Fig. 1, the frames are taken at various physical time steps equal to a multiple of 0.05 ps (which is equal to 10 times the computational time step). We have plotted the logarithmic contour plots ($\log_{10}$) of the probability density $|\psi(t)|^2$. For $k_0 \approx$ 0.0094Å^{-1}, the electron De Broglie wavelength ($\lambda = 2\pi/k_0$) is about 660Å and therefore bigger than the slit width. Consequently, even though part of the wavepacket ($\geq 35\%$) is transmitted on the right-hand side of the wall at time t≥0.25ps, no diffraction lobes are detectable in the transmitted waveform. For an electron with an incident energy of 45 meV (such an electron is far in the tail of the Fermi distribution in the previous example), $k_0 \approx 0.0265$Å^{-1}, which is close to the diffraction condition, $k_0 \approx 2\pi/w$. In that case, as can be seen in Fig. 2, there is actually a diffraction of the electron wavepacket through the slit as indicated by the existence of three distinct lobes in the contour plots of the charge density profiles in the immediate vicinity (on the right) of the slit at time t equal 0.15 ps. However, for later time steps, the diffraction lobes actually smear into a main one which (for this specific case) precludes the actual observation of a Fraunhoffer-like diffraction pattern far from the slit. In fact, the diffraction condition, $k_0 \approx 2\pi/w$, is only met by a small fraction of the various Fourier components of the wavepacket incident on the slit. As a result, the diffraction pattern cannot be as sharp as in the case of an incident plane wave, which is the idealistic situation equivalent to the one used in optics to study light diffraction through a narrow slit. More numerical simulations involving modification of the shape of the wavepacket, potential walls and slit dimensions and extension of the simulation domain need to be performed however before determining if an appropriate set of parameters can eventually lead to a Fraunhoffer-like diffraction pattern far from the slit. The overall shape of the wavepacket behind the slit changes drastically while varying the direction of incident wavevector k_0. This is illustrated in Fig. 3 where the wavepacket with a kinetic energy equal to 10 meV was assumed to be incident at a 45° angle on the slit. This example stresses the importance of collimating the electron beam in order to observe a diffraction pattern with different lobes beyond the slit. This point was already stressed by Kriman and coworkers (1988) in their steady state analysis of the QUADFET.

The width of the potential barrier below the split gate is of primary importance in determining the diffraction pattern behind the split-gate. As a result of the finite width, the electron wavepacket suffer multiple reflections between the potential walls defined by the slit. The resulting spread in transit times introduces additional structure in the diffracted electron beam. For instance, the time frame t = 0.35 ps in Fig. 2 above clearly shows the presence of successive maxima in the transmitted part of the wavepacket in the direction perpendicular to the potential wall. Finally, additional lobes can be clearly seen in the reflected part of the wavepacket, i.e., on the left side of the slit. Such "Back diffraction" would probably also exist while considering realistic split-gate and it could be eliminated by appropriately collimating the incident electron beam.

Conclusions

Preliminary results describing the time-evolution and diffraction of a wavepacket through a narrow slit in a finite potential wall have been reported. In order to save computational time, the size of the slit was chosen to be smaller than the actual split-gate realized with present day technology (Thornton and coworkers (1986); Van Wees and coworkers (1986). Several interesting features in the time evolution of the wavepacket have however been pointed out: (1) A Fresnel-like diffraction pattern is seen at early time once the diffraction requirement is

satisfied, i.e., $k_0 \propto 2\pi/w$ (k_0 being the wavepacket incident wavenumber and w being the slit width); (2) Due to the finite thickness of the slit (in our simulation, the aspect ratio of the slit is equal to 0.5), the main diffracted lobe is found to be fragmented in the direction perpendicular to the slit aperture as a result of the multiple reflections through the slit of finite length; (3) No Fraunhoffer-like diffraction pattern is seen far from the slit even when the slit width is comparable to the electron De Broglie wavelength.

One should emphasize that the Fresnel-like diffraction pattern observed in our simulations is an artifact of the use of infinitely sharp corners in the potential energy profile. A more realistic simulation would require self-consistent calculations of the potential energy profile below a split-gate such as reported by Laux and coworkers (1988). Even with appropriate collimation of a nearly monochromatic incident electron beam, the rounding of the corners below the split-gate on the scale of the wavelength of the incident electron could be the dominant source of suppression of the diffraction pattern at low temperatures.
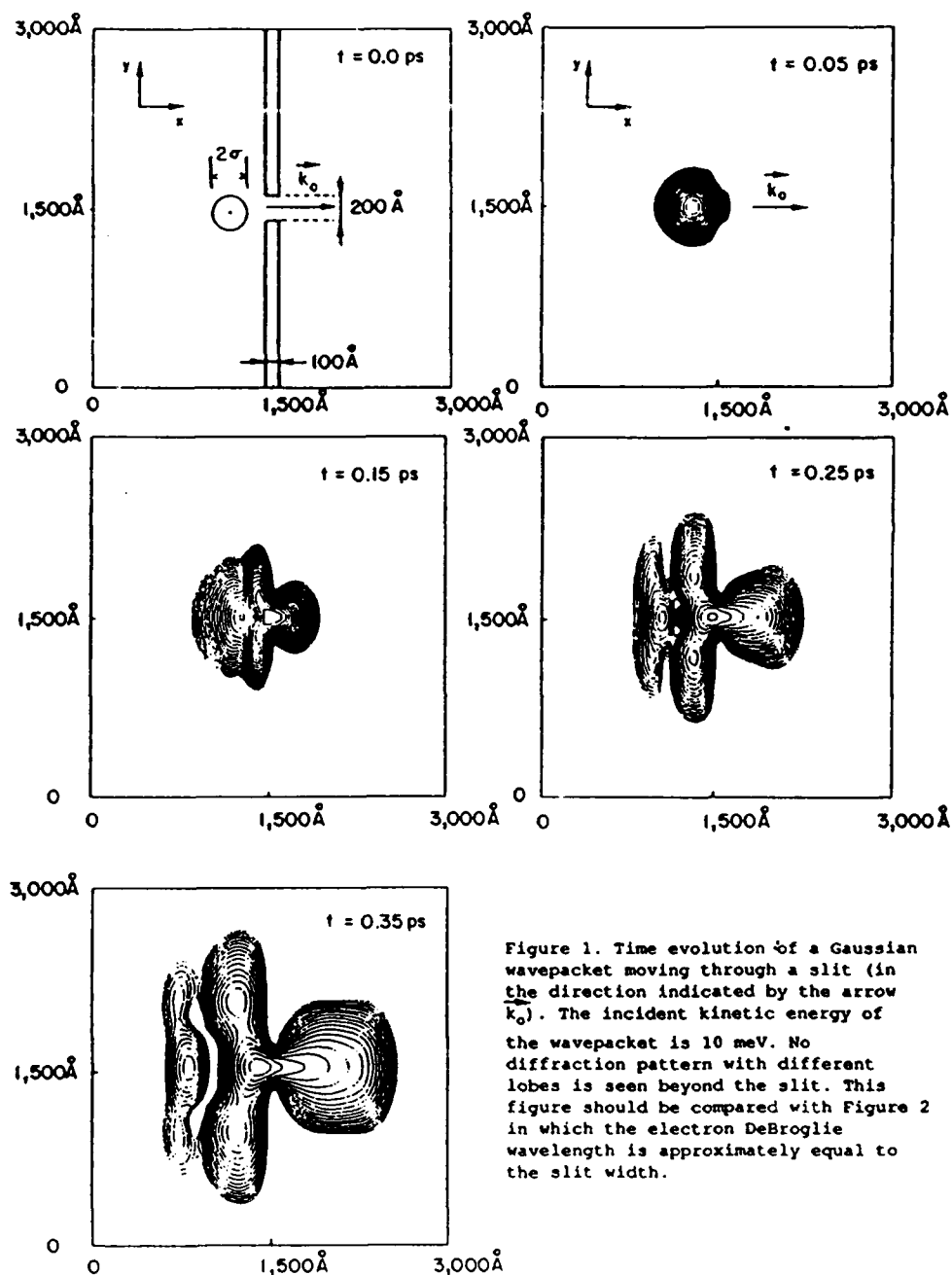

Figure 1. Time evolution of a Gaussian wavepacket moving through a slit (in the direction indicated by the arrow $\vec{k_0}$). The incident kinetic energy of the wavepacket is 10 meV. No diffraction pattern with different lobes is seen beyond the slit. This figure should be compared with Figure 2 in which the electron DeBroglie wavelength is approximately equal to the slit width.
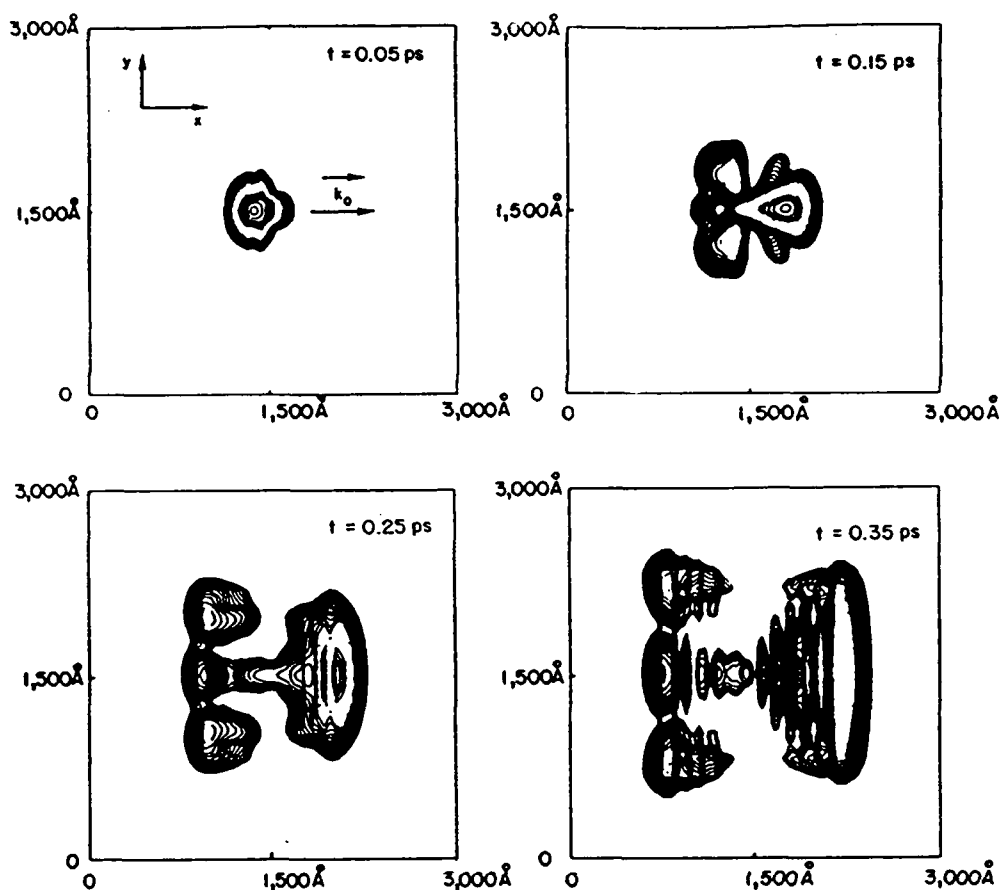
Figure 2. Same as Figure 1 for a wavepacket with an incident energy equal
to 45 meV. Note the formation of diffraction lobes after crossing the slit
which can clearly be seen at time t = 0.15ps. At time t = 0.35ps, the main
transmitted lobe of the diffracted wavepacket is clearly segmented into
various fragments. These are due to multiple reflections of the wavepacket
inside the slit which has finite length. As a result, different parts of the
wavepacket are delayed further in time leading to successive transmitted lobes.
Spurious effect due to reflection at the boundaries of the simulation domain
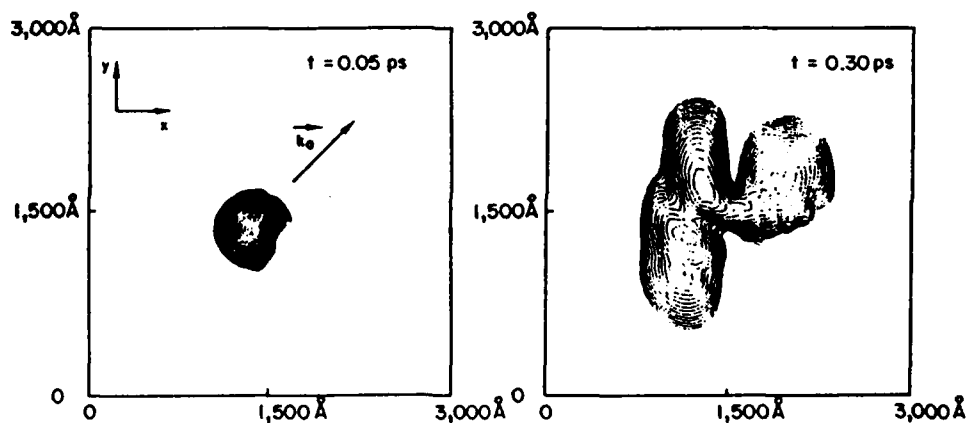are only seen for t = 0.5ps.

Figure 3. Same as Figure 1 while shooting at an angle of 45°
on the slit. The incident kinetic energy is 10 meV.

References

Ancilatto, F., Selloni, A., Xu, L. F., and E. Tosatti, (1989). Phys. Rev. B. 39, 2993.

Bandyopadhyay, S., Bernstein G. H., and W. Porod (1989). Proc. First Conference on Nanostructure Physics and Fabrication, Ed. by M. A. Reed and W. P. Kirk, Academic Press, Boston.

Barker, J. R. (1989). Proc. First Conference on Nanostructure Physics and Fabrication, Ed. by M. A. Reed and W. P. Kirk, Academic Press, Boston; Also, Proc. of San Miniato Workshop on Quantum Transport (1987), San Miniato, Italy.

Cahay, M., Kreskovsky, J. P., and H. L. Grubin (1989). Unpublished.

Douglas, J., and J. E. Gunn (1964). Numer. Math. 6.

Goldberg, A., Shey, H. M. and J. C. Swartz, (1967). Am. J. Phys. 35, 177.

Hauge, E. M. and J. A. Stovneng, (1989). Preprint.

Hirayama, Y., Saku, T. and Y. Horikoshi (1989). Phys. Rev. B. 39, 5535.

Kirczenow, G., (1989). Phys. Rev. Lett. 62, 2993 and references therein.

Kriman, A. M., Bernstein, G. H., Haukness, B. S., and D. K. Ferry (1988). Proc. 4th Intl. Conf. on Superlattices, Microstructures and Microelectronics, Trieste, Italy.

Laux, S. E., Frank, D. J., and F. Stern (1988). Surface Science. 196, 101.

Szafer, A. and A. D. Stone (1989). Phys. Rev. Letters. 62, 300.

Thornton, T. J., Pepper, M., Ahmed, H., Andrews, D., and G. J. Davies (1986) Phys. Rev. Letters 56, 1198.

Van Wees, B. J., Van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwenhoven, L. P., Van der Marel, D., and C. T. Foxon (1988). Phys. Rev. Letters. 60, 848.

Zheng, H. Z., Wei, H. P., Tsui, D. C., and G. Weimann (1986). Phys. Rev. B. 34, 5635.

Acknowledgement

DOUBLED FREQUENCY OF THE CONDUCTANCE MINIMA IN ELECTROSTATIC AHARONOV-BOHM OSCILLATIONS IN ONE-DIMENSIONAL RINGS[1]

M. Cahay[a], S. Bandyopadhyay[b] and H. L. Grubin[a]

[a]Scientific Research Associates, Inc.
Glastonbury, Connecticut 06033

[b]Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

We predict the existence of *two* different sets of conductance minima in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect. The two sets of minima arise from two different conditions and effectively *double* the frequency of the conductance troughs in the oscillations. This makes the frequency of the troughs *twice* that predicted by the Aharonov-Bohm effect. We discuss the origin of this feature along with the effects of temperature and elastic scattering. We also compare it with the magnetostatic Aronov-Al'tshuler-Spivak effect and point out the similarities and differences.

I. INTRODUCTION

Oscillatory conductance due to the electrostatic Aharonov-Bohm effect has been predicted for a variety of ring structures along with potential device applications of that effect. In this paper, we point out an intriguing feature in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect. Unlike in the magnetostatic effect, the conductance in the electrostatic effect reaches its minimum under two *different* conditions which gives rise to *two-*

---

distinct and independent sets of conductance minima in the oscillations. One set of minima arises from the usual destructive interference of transmitted electrons and the other arises from constructive interference of reflected electrons. The minima in each individual set recur in the oscillations with the periodicity predicted by the Aharonov-Bohm effect, but the separation between two adjacent minima (belonging to the two different sets) is smaller than and unrelated to the Aharonov-Bohm periodicity. In the following Sections, we establish this feature and discuss various issues related to it.

II. THEORY

The conductance G of a one-dimensional structure in the linear response regime is given by the two-probe Landauer or Tsu-Esaki formula [1]

$$G = \frac{e^2}{2hkT} \int dE \ |T_{total}(E)|^2 \ sech^2(\frac{E - E_F}{2kT}) , \qquad (1)$$

where $T_{total}(E)$ is the transmission coefficient of an electron with incident energy E through the entire structure (i.e. from one contact to the other), T is the temperature and E_F is the Fermi level.

The problem of calculating the conductance G is essentially the problem of calculating T_{total}. The quantity T_{total} can be found from the overall scattering matrix for the structure. For a ring structure, the overall scattering matrix is determined by cascading three scattering matrices [2] representing propagation from the left lead of the ring to the two interfering paths, propagation along the paths, and propagation from the paths to the right lead. For simplicity, we will represent the first and the last of these scattering matrices by the so-called Shapiro matrix which is defined in Ref. 3.

A. Ballistic Transport

In the case of ballistic transport, cascading the aforementioned three scattering matrices (according to the prescription of Ref. 2) yields the overall scattering matrix and the transmission T_{total} [1,4] as

$$T_{total} = \frac{\epsilon[(t_1 + t_2) - (b - a)^2 t_1 t_2 (t_1' + t_2')]}{[1 - t_1(a^2 t_1' + b^2 t_2')][1 - t_2(a^2 t_2' + b^2 t_1')] - a^2 b^2 t_1 t_2 (t_1' + t_2')^2} \qquad (2)$$

where ϵ, a and b are the elements of the Shapiro matrix[2], and t and r stand for transmission and reflection amplitudes within the two interfering paths. The subscripts '1' and '2' identify the corresponding path and the unprimed and primed quantities are associated with forward and reverse propagation of the electron.

---

[2]For a definition of these elements, see Ref. 1, 3 or 4.

In the presence of an external potential V inducing the electrostatic Aharonov-Bohm effect, t_1, t_2, t_1' and t_2' transform according to the following rule [4]:

$$\begin{pmatrix} t_1 \rightarrow \hat{t}_1 & t_1' \rightarrow \hat{t}_1' \\ t_2 \rightarrow \hat{t}_1 e^{i\phi} & t_2' \rightarrow \hat{t}_1' e^{i\phi} \end{pmatrix} , \qquad (3)$$

where the quantities with the "hats" represent the transmission amplitudes in the absence of the external potential V, and ϕ is the electrostatic Aharonov-Bohm phase-shift between the two paths induced by V and given by

$$\phi = \frac{e}{\hbar} V < \tau_t > = \frac{\sqrt{2m^* E}}{\hbar} [\sqrt{1 + \frac{eV}{E}} - 1] L \qquad (4)$$

Here $< \tau_t >$ is the harmonic mean of the transit times through the two paths which depends on V and the kinetic energy E of the electrons, m^* is the electron's effective mass and L is the length of each path.

Using the transformations given by Equation (3) in Equation (2) and assuming that in the absence of the external potential V the two paths are identical in all respects (i.e. $\hat{t}_1 = \hat{t}_2$ and $\hat{t}_1' = \hat{t}_2'$), we obtain

$$T_{total}(\phi) = \frac{e\hat{t}_1(1 + e^{i\phi})(1 - (b-a)^2 \hat{t}_1 \hat{t}_1' e^{i\phi})}{D(\hat{t}_1, a, b, \phi)} , \qquad (5)$$

where the denominator D is a function of $\hat{t}_1$, a, b and ϕ.

We find from the above equation that $T_{total}(\phi)$ vanishes and hence the conductance (see Equation (1)) reaches a minimum whenever

$$\phi = (2n+1)\pi , \quad i.e. \ when \quad \frac{\sqrt{2m^* E}}{\hbar}[\sqrt{1 + \frac{eV}{E}} - 1]L = (2n+1)\pi$$
$$(6)$$

This gives the usual conductance minima (which we call the *primary* minima) associated with destructive interference of transmitted electrons.

However, we find from Equation (5) that $T_{total}(\phi)$ also vanishes whenever

$$(b-a)^2 \hat{t}_1 \hat{t}_1' e^{i\phi} = 1 \qquad (7)$$

From the unitarity of the Shapiro matrix (see Ref. 4) it can be shown that $b - a$ differs from unity only by a constant phase factor, i.e

$$b - a = e^{i\nu} \qquad (8)$$

Now since in ballistic transport $\hat{t}_1 = \hat{t}'_1 = e^{ikL}$ (where k is the electron's wavevector in either path in the absence of the external potential V), Equation (7) really corresponds to the condition

$$2k_1 L + \phi + 2\nu = \frac{\sqrt{2m^* E}}{\hbar}[\sqrt{1 + \frac{eV}{E}} + 1]L + 2\nu = 2m\pi \quad (9)$$

Whenever Equation (9) is satisfied, another set of conductance minima should appear in the oscillations since the numerator of $T_{total}(\phi)$ goes to zero and the conductance should fall to a minimum unless the denominator of $T_{total}(\phi)$ also happens to go to zero at the same time. It is easy to see that the denominator of $T_{total}(\phi)$ vanishes whenever ϕ is an even multiple of π. Hence, unless Equation (9) is satisfied only by those values of ϕ that are even multiples of π (which requires $2(k_1 L + \nu)$ to be also an even multiple of π), the conductance of the structure should reach a minimum whenever ϕ satisfies Equation (9). This gives rise to an additional set of minima which we call the *secondary minima*. Actually, the secondary minima always occur unless $2(k_1 L + \nu)$ is an even or an odd multiple of π. The latter case is not proved here for the sake of brevity, but is proved in Ref. 4.

B. Diffusive Transport

In the case of diffusive transport, $T_{total}(\phi)$ can again be found from the prescription of Ref. 2, except that now we have to evaluate it numerically. We have calculated the conductance G vs. the electrostatic potential V for both ballistic and diffusive transport. The results are displayed in Fig. 1. The secondary minima are not washed out by elastic scattering in the weak localization regime. However, they begin to wash out with the onset of strong localization and with increasing temperature. The effect of temperature has been discussed in Ref. 4. Note also the interesting feature exhibited by the secondary minima; they become more and more pronounced in the higher cycles of oscillations (increasing V) unlike the primary minima. This implies that in an experimental situation, even if the secondary minima cannot be observed in the first few cycles, they could show up in the later cycles.

III. DISCUSSION

Before concluding this paper, we briefly discuss the origin of the secondary minima. Equation (9), which predicts the existence of the secondary minima in the ballistic case, physically represents the condition that an electron reflected *around* the ring interferes constructively with itself at its point of entry into the ring. This minimizes the conductance by maximizing the reflection. Such a phenomenon can be viewed as some kind of "coherent backscattering", but it is not exactly similar to the magnetostatic Aronov-Al'tshuler-Spivak (AAS) effect which also involves backscattering, but specifically involves interference of two backscattered *time-reversed paths*. Conductance modulation due to the interference of time-reversed
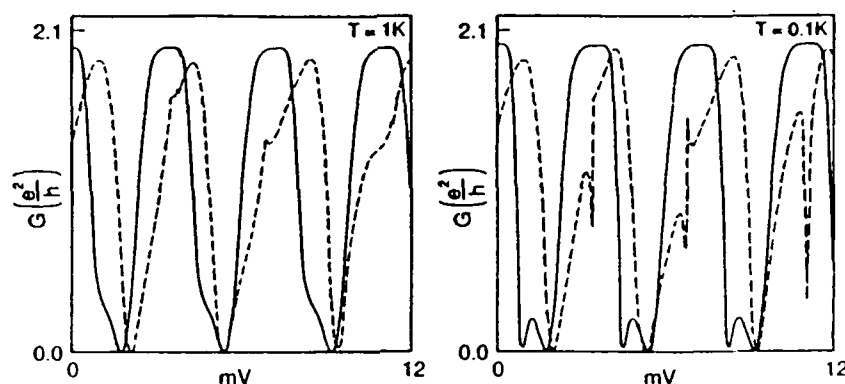
Fig. 1. Electrostatic A-B oscillations in a 1-d ring. The length of each path is 5000 Å. The carrier concentration is 1.55 x 10^6 cm^{-1} and the parameter ε = 0.5. The solid curve is for ballistic transport and the broken curve is for diffusive transport. In the latter case, there are 10 elastic scatterers in each path arbitrarily located. Strong localization would have set in if there were 33 scatterers in either path. In both ballistic and diffusive transport, the secondary minima are bleached out much more rapidly than the primary minima as the temperature is increased.

paths cannot occur in the electrostatic case since the time reversed paths always interfere constructively and an external electrostatic potential cannot change that[3]. However, in spite of this basic difference, there is undeniably the superficial similarity between the two effects in that they both double the frequency of the conductance troughs in the oscillations.

REFERENCES

1. Bandyopadhyay, S., and Porod, W. (1988a). Appl. Phys. Lett. 53, 2323. ; Bandyopadhyay, S., and Porod, W. (1988b). Superlattices and Microstructures, to appear.
2. Cahay, M., McLennan, M., and Datta, S. (1988). Phys. Rev. B15., 37, 10125.
3. Shapiro, B. (1983). Phys. Rev. Lett. 50, 747.
4. Cahay, M., Bandyopadhyay, S., and Grubin, H. L. (1989). to appear Phys. Rev. B15 (Rapid Communication).

---

[3]This happens because the phase-shifts suffered by an electron in traveling along opposite directions have the *same* sign in the electrostatic case, but *opposite* signs in the magnetostatic case.

411